The London School of Economics and Political Science

Vying for Votes:
A Comparison of Off- and Online Election Campaign Strategies

Christian Mueller

A thesis submitted to the Department of Methodology of the London School of Economics and Political Science for the degree of Doctor of Philosophy, London, September 2023

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

I declare that my thesis consists of 30110 words.

Statement of co-authored work

I confirm that Chapter II was jointly co-authored with Dr Florian Foos and Prof Peter Jon and I contributed 40% of this work.

I confirm that Chapter IV was jointly co-authored with Dr Bruno Wüest and Thomas Willi and I contributed 40% of this work.

Statement of use of third party for editorial help

I can confirm that my thesis was copy edited for conventions of language, spelling and grammar by the proofreading service of LSE's Language Centre and the Grammarly web service.

Abstract

Elections are a fundamental part of the political process. Today, election campaigns not only focus on traditional strategies to attract voters but also use social media as a tool. I analyse and compare campaign strategies from three different angles in this thesis. The first paper examines how Get Out The Vote (GOTV) leaflets can influence turnout for the neighbours of households which receive flyers. Focusing on a GOTV campaign during a UK election, I show that spillover effects for party supporters are lower when the share of rival party supporters is high. At the same time, turnout spills over to rival party supporters in mixed partisan neighbourhoods. Turning to online election campaigns, the second paper analyses social media usage from the lens of parties in Switzerland. Using data from party-affiliated Twitter accounts during the 2015 Federal Election, I study how cohesively parties organise their members and how coherent parties' programmatic messaging is. The results show that smaller-sized and newer parties have higher organisational cohesion and that most parties exhibit low levels of programmatic coherence. Switching the lens to candidates, I analyse social media use by candidates during the 2019 European Parliament elections. The third paper introduces a comprehensive dataset of parties, candidates, and their Facebook and Twitter accounts and describes how the data was collated. To show the range of potential applications of the dataset, I outline an analysis of social media adoption and discuss other research areas in which this data could be useful. The final paper studies whether electoral systems guide if and how individual candidates use social media. The results show that when the electoral system favours person- over party-based campaigning, candidates do not use Twitter more but adapt their communication style to engage voters instead of broadcasting information.

To Barbara, Michael and Werner.

It would not have been possible without you.

In loving memory of
Werner (1925–2013) and Trudi (1925–2021)
Stricker-Hunziker.

Acknowledgements

First and foremost, I want to thank my supervisors, Ken Benoit and Jouni Kuha. They supported me every step of the way. I am especially grateful for their help during difficult phases.

I also want to thank my co-authors Florian Foos, Peter John, Kevin Cunningham, Bruno Wüest, and Thomas Willi for collaborating on two projects.

I also want to thank Arthur Spirling and Joshua Tucker, who allowed me to visit New York University. The unique environment and the people at the Center for Data Science and the Social Media and Public Participation lab (now the Center for Social Media and Politics) proved very fruitful for my studies. I gratefully acknowledge financial support from the Swiss National Science Foundation through their doc.mobility fellowship programme (grant no. P1SKP1_181328), which financed this visit.

I am indebted to Kelsey Siegel, who helped me during the final phase of the PhD.

Last but not least, I want to acknowledge a few people who took me in as a BA student at the University of Zurich: Lucas Leeman, who gave me my first job as a research assistant, and Marco Steenbergen and Dominik Hangartner, who trusted me at an early stage to make small contributions to several research projects.

Contents

List of Tables

List of Figures

Chapter I

Introduction

Downs (1957, 30) famously assumes that the primary motivation of politicians is their "desire for power, prestige, and income." To achieve these goals, they must ensure that their party gathers enough votes to be elected to office. This is, of course, a pointed statement in the context of his *Economic Theory of Democracy*, but it highlights the fundamental role of elections as gatekeepers of political influence. All parties must go through the electoral arena and devise campaign strategies to attract voters, whether they are vote-seeking, office-seeking or policy-seeking (Strom 1990).

In this thesis, I present three research projects which examine election campaign strategies from different angles. The first project in Chapter II investigates the effect of Get Out The Vote (GOTV) leaflets—a traditional, off-line strategy to attract voters—from the angle the Huckfeldt and Sprague's (1987; 1995) theory of political mobilisation within social networks. The paper develops a theory of when the effect of the GOTV leaflets should spill over—from households which received a leaflet to other households in the neighbourhood—and tests it on a leaflet campaign conducted in the UK.

The second project in Chapter IV turns to online election campaigns on social media and analyses how politicians used Twitter during the 2015 Federal Elections in Switzerland from the parties' perspective. It focuses on why and how successfully parties try to control their members' organisational cohesion and programmatic coherence.

The third project in Chapters VI and VII switches the focus to candidates and examines their adoption of social media strategies during the 2019 European Parliament elections. I describe the collection of a comprehensive dataset of candidate names and their social media accounts (Chapter VI), which I then use to examine whether electoral system effects influence how often they use Twitter and what communication style they adopt (Chapter VII).

Chapter II

(Paper 1) Social Mobilisation in Partisan Spaces

Abstract

Three decades ago, Huckfeldt and Sprague hypothesised that partisan context constrains information sharing between neighbours. We develop their theory to identify implications for campaign mobilisation in homogeneous and mixed partisan contexts. We argue that Get Out The Vote (GOTV) spillover effects should vary with the proportion of rival party supporters in a neighbourhood. We test this expectation of differential spillover between households that were either included or excluded, pre-random assignment, from a street-level GOTV experiment. We estimate neighbourhood party preferences based on targeting data, made available by the UK Labour Party. We find that GOTV spillover effects for party supporters are smaller in neighbourhoods that include larger shares of rival party supporters. Rival partisans are mobilised in mixed neighbourhoods, where the probability of spillovers from mixed partisan households is higher. This paper extends Huckfeldt and Sprague's theory, and demonstrates the importance of social dynamics for parties' campaign strategies.

This chapter is published as:

Florian Foos, Peter John, Christian Müller and Kevin Cunningham. 2021. "Social Mobilization in Partisan Spaces". *The Journal of Politics* 83, no. 3 (July): 1190–1197.

DOI: 10.1086/710970.

The text is based on the pre-publication manuscript with minor adjustments to improve legibility and clarity for the thesis format.

1 Motivation

As Huckfeldt et al. (1993, 366) wrote over a quarter of a century ago, "the central motivation for contextual theories of politics is the idea that patterns of social interaction are influenced by surrounding population distributions". There has been a lively debate about how political heterogeneity conditions political mobilisation within social networks, such as the household and the neighbourhood (Mutz 2006; Klofstad et al. 2013; Bello and Rolfe 2014). However, little progress has been made in applying Huckfeldt and Sprague's (1987; 1995) insights about how the partisan composition of a social space constrains social and political interactions between citizens, to the effectiveness of parties' ground campaigns. There is a large observational (Whiteley and Seyd 1994) and growing experimental literature (Foos and Rooij 2017b; Townsley 2018) on the effects of partisan campaign canvassing in the United Kingdom that confirms findings from the United States: Get Out The Vote (GOTV) campaigns are effective at increasing turnout (Nickerson et al. 2006).

Moreover, GOTV spillover experiments show that campaign contact spills over within households (Nickerson 2008; Dahlgaard et al. 2021; Foos and Rooij 2017a), giving credibility to the theoretical claim supported by observational research (Fieldhouse and Cutts 2016) that "all turnout is, in a sense, mobilized, with much of the mobilization occurring indirectly" (Rolfe 2012, 121). However, there is mixed causally-identified evidence on GOTV spillover effects between neighbours (Gay 2011; Sinclair et al. 2012). To identify treatment effects, most GOTV experiments assume that spillovers exist within, but not between households. This assumption runs counter to work which suggests that voters' decision whether to turn out is influenced by the behaviour of their peers (Rolfe 2012). This suggests that relatively large spillover effects can materialise due to cascades of mobilisation from neighbour to neighbour (Fowler 2005; Fieldhouse et al. 2015).

Further, there is little theory and evidence on whether dynamics of intra-household mobilisation are linked to inter-household mobilisation within neighbourhoods. In this paper, we provide theory and evidence to fill this gap. Neighbourhoods, in contrast to households, are usually made up of weak ties (Morey et al. 2012), and therefore we follow Huckfeldt and Sprague's (1987) prediction that neighbours intend to share information with co-partisans.

However, even if this is the case, campaign effects can still spill over to supporters of rival parties. Based on a campaign experiment conducted in Birmingham, UK, Foos and Rooij (2017a) show that a party's canvassing campaign directly mobilised targeted voters, but also indirectly mobilised both supporters and opponents of the party that shared the same household. Since we find a higher share of mixed partisan households in mixed-partisan neighbourhoods, we should expect members of mixed partisan households to mobilise neighbours who share the same party preferences. In the worst case for parties' campaigns, these patterns of between-household contagion can increase turnout overall, but render parties' mobilisation campaigns ineffective at moving vote shares.

Using targeting data collected by UK Labour Party canvassers and leveraging the spillover effects of a randomised field experiment, we confirm the existence of such mobilisation dynamics that link mixed partisan households to politically heterogeneous neighbourhoods. The indirect mobilisation effects of a Labour Party leafleting campaign varied among neighbours conditional on the partisan composition of neighbourhoods. In neighbourhoods with a large share of Labour supporters, GOTV effects spilled over to Labour supporters who were not initially targeted. There were no spillover effects to other Labour supporters in neighbourhoods where a majority of residents supported a rival party. Supporters of rival parties were mobilised in mixed partisan neighbourhoods where spillovers could originate from households where Labour supporters live with rival partisans.

2 From Intra- to Inter-Household Mobilisation

The local area of the neighbourhood is an important site of social interaction (Enos 2017). Even though informal and low intensity in character, the neighbourhood can have a strong impact on social and political outcomes over time (Gay 2011). From prior research, we expect that the partisan context in which election campaigns take place should affect the formation and maintenance of discussion networks within neighbourhoods (Huckfeldt and Sprague 1987). If neighbours prefer to share information with like-minded others, GOTV spillover effects between neighbours should be conditional on shared partisanship.

The network literature suggests that the strength of social ties conditions how willing individuals are to engage with others who disagree with them politically (Morey et al. 2012). Household members who support different parties continue to talk politics (Bello and Rolfe 2014) and mobilise each other during election campaigns (Foos and Rooij 2017a). In contrast, when ties are weak, individuals may refrain from sharing information (Mutz 2006). However, the link between household and neighbourhood mobilisation dynamics has rarely been investigated. Parties usually target voters for GOTV who are likely to support them, hence they are unlikely to target households that do not contain at least one pre-identified party supporter. Canvassers are homophilous and are more likely to talk to voters who are similar to themselves (Nall et al. 2017). However, from prior research we know that when canvassers speak to supporters who live in mixed partisan households, rival partisans will be mobilised to vote (Foos and Rooij 2017a). Even if parties correctly identify supporters and opponents based on detailed targeting data, citizens are at an information disadvantage. They should be more likely to mis-identify co-partisans in mixed partisan neighbourhoods (Huckfeldt and Sprague 1987). In both cases, even if the party as well as party supporters intend to exchange information only with co-partisans, targeting mixed neighbourhoods can have the unintended consequence of mobilizing both co-partisans and supporters of rival parties. This logic is displayed in Appendix Figure 10.

Household dynamics have implications for political mobilisation within neighbourhoods because political information flows between citizens depend on the strength of personal ties between supporters of different parties, i.e. whether they share the same household. When Labour canvassers mobilise Labour voters in homogeneous households, the contacted individual not only mobilises her household member, but this contact also spills over to other Labour partisans in neighbouring households. When canvassers target mixed partisan households, partisans of all stripes are mobilised. In this case, even if spillover between households in a neighbourhood flows between co-partisans, rival partisans who are indirectly mobilised within the household can mobilise their co-partisans within the neighbourhood. This is the opposite of the intended effect of a partisan GOTV campaign.

3 Experimental Set-Up

To test these expectations, we use data on individuals excluded pre-random assignment from a partisan GOTV experiment that we previously conducted in collaboration with the UK Labour Party during the 2014 European and local election campaign in one English parliamentary constituency and local government jurisdiction. The constituency has a large Labour majority, and the wards the party chose to campaign in were Labour held.[1] We compare the turnout rates of non-experimental subjects living in streets assigned to treatment with non-experimental subjects living in control streets. We also use geocoding to investigate whether treatment effects vary conditional on whether the members of the closest household support Labour or a rival party. The analysis focuses on the indirect mobilisation effects of a GOTV-leafleting campaign conducted by the UK Labour Party.[2] The treatment in the original experiment was a partisan leaflet that highlighted either the Conservative government's failure on the NHS or on crime and policing, and which was put through the door by local Labour Party volunteers. Besides the issue-specific content, all leaflets included an appeal to vote Labour in the local and European elections on 23 May 2014.[3] The treatment materials are displayed in Appendix Figure 11. As specified in the pre-analysis plan[4], to maximise statistical power for the spillover analysis, we combine both treatment arms into one. Validated turnout was obtained at the individual level from the public register, then merged with the random assignment and pre-treatment covariates.

The original set-up of the randomised field experiment lends itself to the analysis of social influence between neighbours because, initially, large numbers of households were excluded from the experiment based on design and feasibility considerations. By retracing the restrictions which were used to create the original experimental sample, we are able to generate a data set of households located on treatment and control streets which were not part of the experiment. The detailed sample selection procedure is described in Data Ap-

1. See Appendix A.2 for background details on research site.
2. The party campaigned as it normally would, and the experimental assignment to treatment and control streets reflected the need to allocate scarce resources. Even under the most conservative assumptions, the 3% point increase in turnout on treatment streets that we estimate could neither have affected the council majority, nor the seat allocation in the European or local elections.
3. As is well known (Reif and Schmitt 1980, 14), parties campaign on national issues in second order elections because voters consider national politics to be more important than supranational and/or local matters.
4. See http://egap.org/registration-details/4388 for re-registered hypotheses and the de-identified PAP.

pendix B.1. This enables us to use two strategies to identify spillover effects between neighbours. First, we operationalise the larger neighbourhood of each subject as the segment of the street the subject lives on, which falls into a single electoral ward. We then compare individuals living in households excluded from the original experiment located in treatment street segments to individuals living in excluded households located in control street segments, under the identification assumption that spillovers occur within but not between street segments. Second, we define the immediate neighbours of each individual as those subjects living in the most proximate household on the same street segment. We first locate all the household addresses on the map through geocoding and we compute Euclidean—or direct line—distances between households that are located on the same street segment. For each non-experimental household, we then identify the closest experimental household by minimising this distance. If there are ties, we average across all equally proximate households. The geolocation and distance computation is explained in detail in Data Appendix B.2. The resulting dataset contains individual-level information on the neighbours of 16,014 non-experimental subjects living on 615 street segments.

4 Partisan Heterogeneity

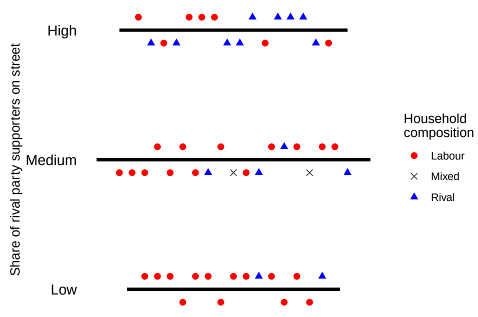


Figure 1: Partisan heterogeneity within and between streets.

As pre-registered, we define partisan heterogeneity as the share of rival party supporters who live on the same street segment (analysis 1), and the share of rival party supporters who live in the most proximate household (analysis 2). There are, of course, many reasonable alternative ways of how we could define partisan heterogeneity (e.g. the proportion of Labour supporters), and how we could define a neighbourhood (e.g. by zip-code), and this is the reason for why we registered our preferred definition in the pre-analysis plan. Our dataset contains information about the party preferences of 8,375 experimental subjects and 16,014 non-experimental subjects. The data on party preferences used to estimate the share of rival party supporters per street segment and most proximate household is estimated based on pre-treatment targeting data collected by Labour Party canvassers. For an extensive validation of the canvassing-based voting intention measure used in this paper, see Foos (2018) who uses the same measurement instrument in a different constituency. The shares of party self-identifiers in the experimental and the non-experimental samples are displayed in Appendix Figure 15. Moreover, the distribution of rival party supporters per street is displayed in Appendix Figure 12, which shows that the share of rival partisans in a neighbourhood is approximately normally distributed in our sample. Figure 13 in the Appendix shows the correlation between the partisan composition of neighbourhoods and the partisan composition of the most proximate household, the two measures used in this paper. Mixed partisan neighbourhoods have a significantly larger share of mixed partisan households. That means that in line with figure 1, individuals who live in mixed partisan neighbourhoods are also more likely to live next to a mixed partisan experimental household.

5 Analysis

We estimate the following linear models, clustering standard errors at the level of assignment, the street level:

$$Y_i = \alpha + \beta Z_i + \epsilon_i \text{ (1)}$$

$$Y_j = \alpha + \beta Z_j + \epsilon_i \text{ (2)}$$

$$Y_j = \alpha + \beta_1 Z_j + \beta_2 X_{1ij} + \beta_3 X_{1ij} * Z_j + \epsilon_{ij} \text{ (3)},$$

where Y_i is validated individual-level turnout (1 or 0) for subjects living in households originally included in the experiment, Y_j is validated individual-level turnout for subjects living in households originally excluded from the experiment, α is the turnout rate in the experimental or non-experimental control group, Z is location on a treatment (1) or control (0) street, X_1 is the share of rival party households in a street, and ϵ is the error term. All models also include fixed effects for experimental blocks (electoral wards).

We also pre-registered the following equation, which identifies indirect mobilisation effects conditional on the partisan composition of the most proximate household:

$$Y_j = \alpha + \beta_1 Z_j + \beta_2 X_{2ij} + \beta_3 X_{2ij} * Z_j + \beta_4 X_{3ij} + \beta_5 X_{3ij} * Z_j + \beta_6 X_{2ij} * X_{3ij} + \beta_7 X_{2ij} * X_{3ij} * Z_j + \epsilon_{ij},$$

where X_2 is the share of rival party supporters within the most proximate household, and X_3 is the Euclidean distance to the closest household. Since linear interaction terms lack common support, we diverge from our PAP and estimate the interaction effects using the binning method proposed by Hainmueller et al. (2019). We now estimate the following equation in Table 18:

$$Y_j = \alpha + \beta_1 Z_j + \beta_2 X_{2ij} + \beta_3 X_{3ij} + \beta_4 X_{2ij} * Z_j + \beta_5 X_{3ij} * Z_j + \beta_6 \mathrm{X}'_{ij} + \epsilon_{ij} \quad (4),$$

where X_2 are mixed partisan households, X_3 are rival-only households, and X'_{ij} is a matrix of k pre-treatment covariates for n subjects in household j. Pre-treatment covariates include the Euclidean distance to the closest experimental household, the number of subjects in the closest experimental household, and the interaction between the number of subjects in the closest experimental household and treatment assignment. Moreover, we further diverge from our pre-analysis plan by restricting our sample to subjects for whom the Labour Party collected pre-treatment data on party preferences. We did not anticipate that these data would be missing for around 50% of our sample. We present the main analysis for the complete sample including those subjects that do not identify with any party in Figure 16. Finally, we report the results of the analyses conditioning on the partisan composition of the neighbourhood and the partisan composition of the household as pre-specified for the full sample of party supporters, and separately for Labour and rival party supporters. As pre-specified we report both unadjusted and covariate-adjusted ITTs. We report the covariate-adjusted analyses (turnout in the 2013 local election, household size, and gender) in Appendix Tables 18 and 20.

6 Results

We conduct differential attrition checks and balance checks using randomization inference. The p-value of .39 indicates that there is no evidence of differential attrition as a function of treatment assignment (for a full explanation of the procedure see Figure 14 in the Appendix). Table 16 shows balance on available pre-treatment covariates, household size, gender, turnout in the 2013 local elections, as well as Labour or rival party identification. We also conduct a balance test using randomization-inference, which shows that in 875 of 5000 simulated random assignments, imbalances between treatment and control groups were larger or as large as in our dataset, which corresponds to a two-tailed p-value of .18 (see Figure 15 in the Appendix).

First, Table 1 shows that the leaflets successfully mobilised voters to turn out. Experimental subjects on streets assigned to treatment were 2.8 percentage points more likely to turn out than subjects on streets assigned to control. Non-experimental subjects on streets assigned to treatment were around 2.4 percentage points more likely to vote, which indicates that around 86% of the direct effect spilled over. The magnitude of this effect is consistent with cascade effects within neighbourhoods (Fowler 2005), and might also be a function of the relatively low baseline turnout rate (Dahlgaard et al. 2021).

Table 1: ITT of leaflet on turnout of experimental and non-experimental households

	Direct effects	Indirect effects
Control mean	0.489	0.336
	(0.026)	(0.020)
Leaflet	0.028	0.024
	(0.018)	(0.013)
Block fixed effects	Yes	Yes
Cluster standard errors	Yes	Yes
N individual	8375	16014
N cluster	615	615

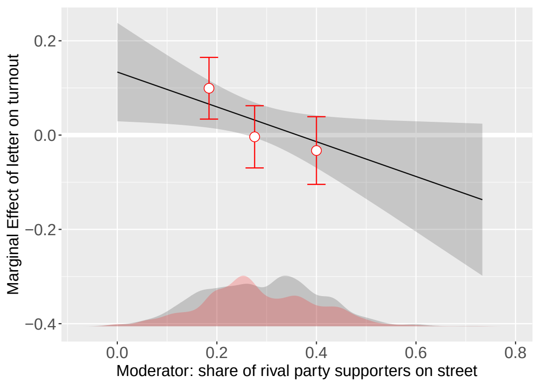
Table 2 displays the indirect Intent-to-Treat (ITT) effects of the Labour GOTV leaflet on validated turnout among individuals living in households that were initially excluded from the experiment, first for all party supporters (columns I and II), and then separately for identified Labour supporters (columns III and IV), and identified supporters of rival parties (columns V and VI). Columns I, III, and V display the main effects, and columns II, IV, and VI introduce the interactions between the treatment and the pre-treatment share of rival party supporters identified to live on the same street segment. This is a treatment-by-covariate interaction, which is not causally identified (Gerber and Green 2012), meaning that we cannot be sure that Conditional Average Treatment Effects arise *because* of the share of rival party supporters. The pre-treatment covariate could be correlated with other unobserved street-level confounders. We account for two of these alternative street segment covariates, the share of experimental subjects per street segment, and the share of subjects per street segment who turned out to vote in the preceding 2013 local elections in Table 17 in Appendix A.9. We also include the interaction of these street section covariates and treatment assignment.

Table 2: ITT of leaflet on turnout of non-experimental subjects conditional on partisan composition

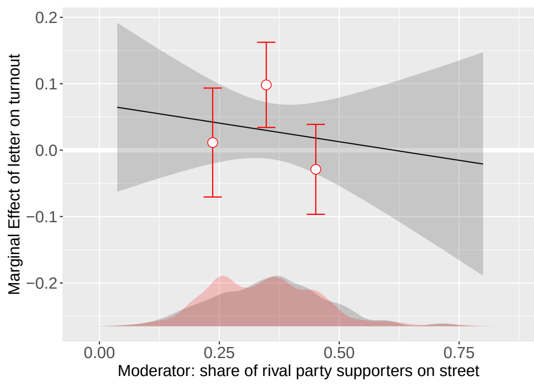
	All Party Supporters		Labour Party Supporters		Rival Party Supporters	
	I	II	III	IV	V	VI
Control mean	0.401	0.296	0.420	0.268	0.389	0.291
	(0.025)	(0.042)	(0.034)	(0.054)	(0.028)	(0.066)
Leaflet	0.027	0.097	0.027	0.134	0.027	0.069
	(0.017)	(0.045)	(0.021)	(0.055)	(0.022)	(0.075)
prop street rival partisan		0.296		0.467		0.256
		(0.111)		(0.153)		(0.169)
Leaflet × prop street rival partisan		−0.215		−0.369		−0.112
		(0.132)		(0.182)		(0.197)
Block fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Cluster standard errors	Yes	Yes	Yes	Yes	Yes	Yes
N individual	10231	10231	5606	5606	4625	4625
N cluster	615	615	615	615	615	615

Note: Standard errors clustered at the street level (CR2). Inverse probability weights for differential probabilities of assignment to treatment between experimental blocks.

Table 2 and Table 17, as well as Figure 16 in the Appendix consistently show that the higher the share of rival party supporters in a neighbourhood, the lower the effects of the GOTV leaflet, which is in line with our expectations of how information sharing between partisans in neighbourhoods should translate into campaign mobilisation. The bins reflect relatively low, medium and high shares of rival party supporters who reside on the same street. We then estimates the ITT of the leaflet within each bin separately. For Labour supporters, spillover effects are positive and significantly different from zero if they reside in predominantly Labour areas. However, the treatment effects are no longer significant once the share of rival party supporters passes 30% of all neighbours. In contrast, spillover effects are estimated to be zero for rival party supporters who reside in neighbourhoods dominated by either party. They only materialise in neighbourhoods that have a mix of Labour and rival party supporters.[5] Figure 2 and Appendix Figure 18 plot the interaction between the treatment and the share of rival party supporters in the neighbourhood separately for Labour party supporters and rival party supporters using the method proposed by Hainmueller et al. (2019). We divide each sample into three equally sized bins, that means there is approximately an equal number of individuals in each bin.



a) Labour supporters



b) Rival party supporters

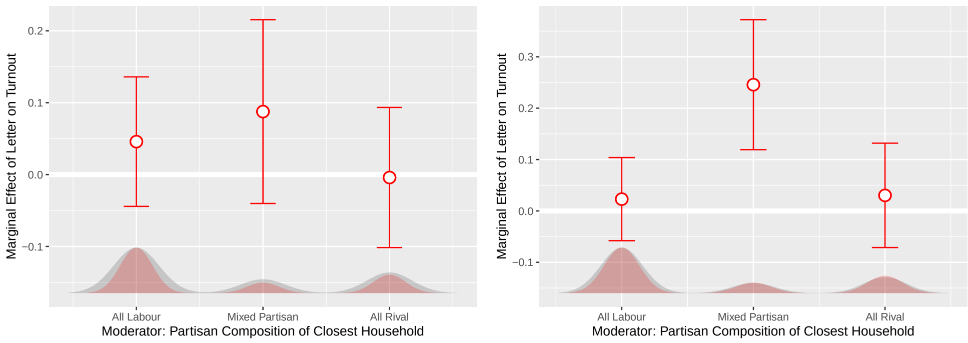
Figure 2: Marginal effects of leaflet conditional on partisan composition of the neighbourhood

Partisan composition of closest household

Having shown that indirect mobilisation effects from a GOTV campaign vary with the partisan composition of the street on which a voter lives, we now consider whether they also

5. We report smooth estimates of the conditional spillover effects and the raw data for treatment and control groups, as well as linear and loess estimates between the moderator and the outcome in Figure 19 of the Appendix. The smooth estimates show no meaningful deviations from the binning estimates reported in the main analysis.

vary if we look at the partisan preferences of the closest neighbouring household. Figure 2 plots the marginal effect of the linear interaction between the treatment and the share of rival party supporters in the most proximate household on the turnout of subjects living in households excluded from the experimental sample. Figure 3 shows that the estimate is zero when the most proximate experimental household only includes rival party supporters. The estimate is 5 percentage-points when the most proximate household consists of Labour Party supporters only, and it is 9 percentage-points when the most proximate household is mixed. Both estimates are not significantly different from zero at the 0.05 level. In contrast, Figure 3 shows that rival party supporters mobilise if the closest experimental household contains a mix of party supporters. Spillover effects from experimental households that contain only Labour party supporters or only rival party supporters are zero.



a) Labour partisans

b) Rival party supporters

Figure 3: Marginal effects of leaflet conditional on partisan composition of the closest household

7 Discussion and Conclusion

This paper shows that campaign effects can spill over within neighbourhoods, and that the magnitude of indirect, social campaign mobilisation is predicted by the partisan composition of neighbourhoods. We use a combination of experimental and targeting data to test the prediction that indirect mobilisation between neighbours should be less likely to occur in neighbourhoods that contain a large share of citizens who oppose the party that initiates contact. In neighbourhoods dominated by a rival party, we would not expect individuals sympathetic to the party that initiates contact to share information with their neighbours who are likely rival party supporters (Huckfeldt and Sprague 1992). In the case of households composed of rival party supporters, we would expect citizens to discount information that comes from a party they do not support (Foos and Rooij 2017b). Our results are consistent with these expectations.

From a theoretical perspective, we link mobilisation dynamics within neighbourhoods to inter-household spillovers between supporters of different parties. If campaign messages are shared in households irrespective of voters' party preferences, then partisan contexts that contain a large enough share of citizens sympathetic to the party that initiates contact should facilitate indirect campaign mobilisation, part of the cascade of mobilisation outlined in other studies (Fowler 2005; Fieldhouse et al. 2015). Our results suggest that this is the case both for in-party supporters and supporters of rival parties. The latter are mobilised in mixed partisan neighbourhoods because the share of mixed partisan households that contain one rival party supporter who can pass on the information, is higher. Understanding the interaction between these intra- and inter-household spillovers in different partisan contexts contributes to the success or failure of a party's GOTV strategy. The results of this paper match the intuition common among canvassers that targeting mixed neighbourhoods can trigger an unintended chain of mobilisation among supporters of rival parties. These unintended consequences may be less likely to occur in contexts that are subject to increasing partisan geographical sorting (Brown and Enos 2021; Martin and Webster 2018). Our findings confirm that heterogeneous social settings are not necessarily prohibitive to facilitating partisan (counter-)mobilisation, especially if they contain a large enough share of citizens who support the party that initiates contact. This finding is consistent with similar findings by Enos (2016) who shows that white Americans are more likely to turn out if living in the direct vicinity of African Americans.

Finally, this study's results have implications for the design of GOTV experiments, which should account for the potential of treatment contagion between neighbouring households. If campaign leaflets or door-to-door canvassing spill over between households, then conducting random assignment at the household level might bias the treatment effect estimator downwards. Our study is a first attempt to integrate randomised campaign experiments and spatial analysis in order to make sense of how social influence operates within political contexts. We hope that it will encourage more sophisticated work at this methodological intersection.

Chapter III

From Off- to Online Campaigning

The previous chapter shows that the turnout effect of GOTV leaflets can spill over to households which were not directly targeted. However, the effect size depends on the distribution of party preferences in the neighbourhood. Spillover effects for co-partisans decrease when the share of rival party supporters increases, while rival partisans are co-mobilised from contact with mixed households in mixed-party neighbourhoods. As discussed above, a large body of literature has established that leaflet campaigns successfully increase turnout (e.g. Nickerson et al. 2006). However, running such campaigns is usually expensive because of the cost of creating, printing and distributing physical leaflets and because they require detailed information about the electorate. Running cost-effective leaflet campaigns involves a database of potential voters to maximise the potential effect. Moreover, one consequence for campaign organisers from the previous chapter is that they also need information on rival partisans to ensure that the effect of the leaflet spills over to co-partisans only.

In contrast, online campaign strategies on social media are often viewed as relatively cheap and generally have low entry barriers. However, they can come with other costs, for example, incivility (Theocharis et al. 2016). However, social media campaigns are not free. Creating a Facebook or Twitter account may be very easy, but it is, at best, the start of a successful social media campaign, as the next chapter shows. After encouraging their politicians to register with Twitter, parties face the challenge of organising them and ensuring that they send a coherent message. The analysis of the 2015 Federal Elections in Switzerland in the next chapter uses a novel, network-based modelling approach to show that smaller and newer parties more successfully maintain organisation cohesion while not staying on message, or programmatic coherence, is generally low.

Chapter IV

(Paper 2) Controlled Networking: Organizational Cohesion and Programmatic Coherence of Swiss Parties on Twitter

Abstract

Political parties are under increasing pressure to extend and activate their voter bases by employing more innovative communication strategies. This article focuses on the social media platform Twitter to explore how well Swiss parties performed in terms of employing digital communication during the 2015 Federal Election Campaign. As such, it uses the follower network as an indicator of organizational cohesion, along with two indicators for programmatic coherence based on Twitter message content. Computing density and centrality statistics allow for the quantification of these two aspects in the party networks, while the non-parametric bootstrap introduces uncertainty of the account sampling process into the analysis. Our results suggest that smaller and newer parties, as well as the Social Democrats, tend to exhibit disproportionally high levels of organizational cohesion. At the same time, most parties show comparable—and also disproportionately low—levels of programmatic coherence compared to those displayed by the Social Democrats.

This chapter is published as:

Bruno Wüest, Christian Mueller and Thomas Willi. 2021. "Controlled networking: Organizational cohesion and programmatic coherence of Swiss parties on Twitter". *Party Politics* 27, no. 3 (May): 581–596. DOI: 10.1177/1354068819872887.

The text is based on the pre-publication manuscript. I made minor adjustments to improve legibility and clarity for the thesis format.

8 Introduction

In today's established democracies, the erosion of party alignments and the growing mediatization of politics have yielded innovations in parties' organizational and programmatic capacities that have become even more essential in election campaigns (Dalton et al. 2000; Strömbäck 2008). A crucial innovation potential thereby lies in the deployment of new communication technologies, such as social media, to reach voters. Accordingly, social media platforms have become regularly used tools in election campaigns across established democracies (see, e.g., Gibson 2015; Vergeer and Hermans 2013; Stromer-Galley 2014; Koc-Michalska et al. 2014; Larsson and Moe 2012; Grant et al. 2010; Theocharis et al. 2016).

As most previous research has focused on the intensity and content of a single politician's social media communication, this study explores social media use at the level of political parties. More specifically, we investigate the performance of Swiss parties on the microblogging platform Twitter, during the 2015 federal election campaign, by assessing a comprehensive set of possible Twitter interactions and dynamics—follower-relations as an indicator of organizational cohesion; re-tweets and text similarity as proxies for programmatic coherence.[6] This approach enables us to link the analysis to established theories on party unity in election campaigns (Jungherr and Theocharis 2017).

Switzerland is a valuable comparative case for exploring the capacity of political parties to mount effective election campaigns via Twitter. First, due to the distinctive consensus-oriented character of its political system, Switzerland has very low barriers limiting entry to the public debate (Höglinger 2008). Relative to most other countries, a larger variety of parties can intensively engage in election campaigns. Second, election campaign managers confirm the increasing importance of social media use (Fichter and Kohler 2015). That said, the size of the Twittersphere in Switzerland is relatively moderate, which makes it much easier to get a comprehensive and comparable sample of Twitter accounts than it would be in larger countries, where the number of users can easily exceed several million. More precisely, our analysis relies on the networking and communication patterns of 1,341 Twitter

6. We use the term *coherence* in connection with the programmatic aspects of the campaign. In other words, programmatic coherence designates largely contradiction-free Twitter communication by a specific party. The term *cohesion*, in contrast, is used to refer to the organizational aspects of a campaign in the sense of tight and dense network structures. We thank an anonymous reviewer for drawing attention to this clarification.

accounts, which have been subjected to bootstrapped network analyses.

Twitter is a popular micro-blogging service in Switzerland. During the 2015 federal elections, about 10% of Switzerland's adult population used Twitter to become informed (NET-Metrix 2015). Moreover, previous research on other countries has established that the political elite and journalists, i.e. actors with disproportionate influence on the campaign, rely even more on micro-blogging than regular citizens do (Wallsten 2010; Himelboim et al. 2013). Switzerland still lags behind; however, during the 2015 elections, about 50% of the 246 members of Switzerland's Federal Assembly and three out of seven Federal Councilors were active on Twitter.[7] Since then, however, more politicians have joined Twitter and communicating on social media has become more important (Fichter and Kohler 2015), which is why we can expect social media to become even more influential as a complement to more classical campaign strategies.

We use Twitter in this study, first, because it is the main social media platform for Swiss politics other than Facebook. Although Facebook has a much larger user base (Freelon 2017), only its pages and not its more relevant personal profiles, are openly accessible. By contrast, information on Twitter accounts is public by default and thereby easily accessible for scientific research (Vergeer 2015). Also, unlike Facebook, Twitter allows non-reciprocal relationships between users, which enables the study of both interactive- and broadcast-style communication. Taken together, Twitter's advantages allow us to conduct a much more thorough analysis of the campaign than would be possible using Facebook data.[8]

Yet this study's contribution is not restricted to Twitter; more generally, it also aims to extend existing research on election campaigns with respect to two important aspects. First, Twitter data permit us to distinguish the more structural features of party organization, such as direct relationships between party-affiliated accounts, from more programmatic ones, such as the congruence in communicative efforts between these accounts. Second, national, regional and local politicians of varying political functions, have been integrated into the analysis, which enables us to study party unity on a much broader set of politicians than

7. According to estimations of party officials, the mean share of active party members on Twitter is slightly lower (Fichter and Kohler 2015).

8. Facebook allows reciprocal relationships between pages, but not between profiles, which would be the equivalent of Twitter accounts.

previous research. Such a research strategy therefore seems generally promising for studying party unity in other strongly federalist states such as Germany or Belgium, where subnational party factions are comparatively powerful.

Our estimates provide evidence that Switzerland's political Twittersphere is shaped by an ideological divide as well as a language-specific separation. Furthermore, it is remarkably representative of 'offline' politics, with the notable exception of the Swiss People's Party (SVP), which is heavily underrepresented on Twitter. The results also underline the importance of Switzerland's federal structure, because it appears to be generally difficult for the national party leadership to control politicians campaigning at the cantonal and local levels. More precisely, in terms of programmatic and organizational unity, networks are denser and more strongly hierarchical at the national level. The analyses of organizational cohesion also show that larger parties are more hierarchically structured than smaller ones. Smaller parties, by contrast, exhibit higher levels of programmatic coherence, in terms of the similarity of their Twitter messages. Finally, Social Democrats stand out for their particularly high levels of organizational and programmatic unity in their networks.

9 Political Campaigning in the Digital Age

Political parties fulfill crucial hinge functions in democratic processes, by activating the electorate and representing specific interests (Gibson et al. 1983). However, three longer-term processes constrain the electoral and representative functions in established democracies. First, Mair (2008) and others convincingly argue that modern governments must increasingly abide by external constraints, to the detriment of their responsiveness to voters. This is especially true for European countries, where governments are simultaneously pressured by European and international policy and market prerogatives (Pontusson and Raess 2012). As Hellwig and Samuels (2007) show, the declining leeway for national policy making leads to a thinning-out of the electoral linkage that makes it harder for citizens to hold governments accountable for policy outcomes. Second, political parties are challenged by the erosion of their organizational basis due to the de-alignment of citizens from previously stable party identifications (Dalton et al. 2000). Because votes have become less structured by party loyalties over the last decades, electoral volatility has increased considerably. Finally, parties are under growing pressure to adapt their campaigning styles to the prerogatives of the mass media (Thesen 2014). This requires parties to streamline their communication strategies in accordance with the relative news value of the messages; the implications of this shift privilege contentious, sensational, personalized, and simplified messages over substantial discussions of policy positions (Esser and Matthes 2013).

Although it is still open to debate how, exactly, international and European constraints, the weakening of party alignments and mediatization have transformed party politics over the last decades, it is clear that these factors have raised the general level of competition in electoral contests. Parties striving for votes and office must increasingly take the center stage of mass-mediated politics to attract the electorate (Kriesi and Trechsel 2008). Because the electorate has become more skeptical and independent over the last few decades, parties must employ innovative campaign strategies. Accordingly, political communication researchers postulate the emergence of distinct changes in the style of campaigning (see Norris 2000). On the one hand, campaign activities are increasingly marked by pressures aimed at professionalization, which implies a strengthening of the parties' leadership and a growing importance of media and public relations consultants (Gibson and Römmele 2001; Bowler et al. 1999). On the other hand, political campaigning is being re-shaped by the continuous integration of new communication technologies (Vergeer et al. 2013). Hence, the ability to deploy new technologies, such as social media, has become a defining element of political campaigns (Esser and Matthes 2013; Obholzer and Daniel 2016).

The deployment of Twitter in election campaigns can be perceived in this context of technological innovations for political campaigning. Twitter is a micro-blogging service with a user community that has been growing since 2006. The service allows users to connect easily to other users (*following*) and to rapidly disseminate and share short messages of 280 characters[9] (*tweets* and *re-tweets*). Users can be identified by their @-*mention* (e.g. @alainberset, the most active Swiss Federal Councilor on Twitter), and trending topics are traceable by their *hashtags* (e.g. #WahlCH15 for the last national elections in Switzerland). There is a burgeoning literature that engages more substantively with the campaigning of political parties

9. In our research period, the maximum number of characters allowed was still 140.

on Twitter. Most of this research focuses on national or regional elections (e.g. Obholzer and Daniel 2016; Graham et al. 2013; Golbeck et al. 2010; Theocharis et al. 2016; Stromer-Galley 2014) and does not apply a network perspective. We extend the scope of our analysis to the municipal level and to politicians active outside the parliamentary arena. Previous studies based on network analyses have focused on single aspects, such as re-tweets or @-mentions (Hemsley et al. 2018; Ausserhofer and Maireder 2013).

Communication via Twitter is cheap, relative to conventional electioneering tools like leaflets, advertisements, and street events. These low costs, paired with the interactive nature of the application, might also enable parties to sustain government responsibility while being responsive, at least remotely, toward their constituencies (Obholzer and Daniel 2016). Accordingly, early research ascribed to Twitter the capacity to engage with constituents, thereby opening the door to more citizen participation in the political process (Theocharis et al. 2016). However, most research indicates that parties' communication on Twitter rarely ever lives up to this normative expectation (Larsson and Moe 2012). Most politicians use the platform in a broadcasting style (Graham et al. 2013). On the one hand, such widespread non-interactivity (or non-responsiveness) by politicians on social media is a reaction to the widespread incivility of user comments (Gervais 2015; Theocharis et al. 2016). On the other hand, non-interactivity clearly serves the individual politicians' interests (Grant et al. 2010). Twitter allows parties to convey political information to their members, journalists, and the broader public rapidly and directly, without bypassing the gatekeepers of traditional mass media (Graham et al. 2013).

Previous work on Twitter suggests that politicians in candidate-centered contests in countries such as the United States and the United Kingdom generally avoid interacting with each other on this medium (Hemsley et al. 2018), instead preferring to use Twitter in a way that Stromer-Galley (2014) calls *controlled interactivity*. Instead of using the full capabilities of Twitter to engage in a genuinely open deliberation with the public, candidates strategically craft their messages to create echo chambers in which the political orientations of their voters are reaffirmed (Colleoni et al. 2014; Vergeer 2015; Conover et al. 2012).

Stromer-Galley's (2014) theory of controlled interactivity was developed to study US presidential election campaigns. These contests are highly candidate-centered and allow parties to play only the subordinate role as election platforms. Analyzing election campaigns in a multi-party system, as Switzerland exemplifies, is different. Political parties have decisive influence over the course of the contest, and previous studies have shown that there is a lot of interaction among the political elites, while users from the broader public are stranded on the network's periphery. This has been found for Denmark, Austria and Switzerland (Larsson and Moe 2013; Ausserhofer and Maireder 2013; Wueest and Mueller 2014). In multi-party electoral contests, the control of the message begins with the parties' control over their own members' communications. To wit, in party-centered contests, parties must engage in controlled networking as the first step of their social media strategy, before single politicians engage in controlled interactivity with the public.

10 Party Unity in Election Campaigns on Twitter

Controlled networking on Twitter is, by no means, a self-propelling process. While campaigning through traditional media channels has become increasingly professionalized, more intense social media use has, ironically, reintroduced amateurism to the political process (Vergeer 2015). Because each single politician has an individual Twitter account, communication is highly decentralized, which gives single candidates the opportunity "to engage in personal promotion outside the auspices of their parties" (Theocharis et al. 2016, 1009). Hence, it would require a considerable effort from parties to control the agenda within their own ranks and to quell messages from dissenters challenging their official campaign programs. However, as outlined further below, asserting and maintaining complete control of the network might not always be a reasonable social media strategy.

Party unity[10] is most commonly studied in parliaments, using different vote- or survey-based measures (Carey and Shugart 1995; Bailer 2017; Hix 2004; Coman 2015; Thorlakson 2009). Classic studies of party unity, by contrast, emphasize the notion that party unity results from several foundational factors, including the centralization of internal decision-making processes, the amount of resources available, the degree of professionalization and the level of programmatic activity (e.g. Gibson et al. 1983). For this analysis, we maintain that

10. Also denoted as the "discipline", "cohesion", "strength", or "integration" of parties (see Verge and Gómez 2012).

this traditional concept of party unity can be summarized along two dimensions. On the one hand, a party can show organizational unity by assuming a tightly-coordinated structure. In such parties, the leadership is able to establish a high level of discipline via centralized organization and comprehensive control over the membership of its members (Carey 2007). On the other hand, a party can exhibit unity via preferences, manifest through the substantively cohesive quality of its members' public campaign messages, or in the congruence of policy decisions (W. L. Benoit et al. 2011). Long before the emergence of social media, "control of 'the message'—the thematically unified collection of issues, frames, talking points, concepts, and images" (Freelon 2017, 170) has been a key campaign objective.

Studies focusing on parliamentary voting records generally struggle to distinguish between these *organizational* and *programmatic* pathways to party unity, as legislative party unity can originate in party discipline, as well as from the cohesive preferences of representatives (see Volden and Bergman 2006). Twitter data, however, facilitates at least some disentanglement of these two aspects. To examine the organizational dimension, we can tap into the follower networks established by Twitter users. Generally, these follower relationships are quite stable as users must add other accounts to their personal networks only once and they rarely ever terminate these connections. To study organizational capacity, we can therefore measure the degree of coordination among accounts affiliated to the same party. Communication on Twitter, by contrast, is highly dynamic, which means that aspects such as the number of re-tweets and @-mentions, as well as the similarity of the tweets' content, indicate coherence among accounts from the same party. Essentially, this means that we can consider each tweet as an instance of programmatic activity, analogous to an election leaflet (Bowler et al. 1999).

Theoretically, a party's communication on Twitter is fragmented into as many accounts as there are party-affiliated accounts on Twitter, which represents, in our case, several hundred accounts for each of the five major Swiss parties. In reality, parties are divided into more or less coherent sub-coalitions, characterized by constant disagreement over goals and the means to pursue them (Kitschelt 1989). One divide is particularly pronounced in the case of Switzerland. Switzerland is a fragmented country where regional (cantonal), and even local, governments exercise a considerable degree of fiscal power and jurisdiction across a wide range of policy areas (Kriesi and Trechsel 2008).

This encourages party organization at the subnational level (Carey 2007). As Thorlakson (2009) has established, heavily decentralized federations with low coordination requirements between federal and state-level governments are likely to be accompanied by highly autonomous state parties. Federal elections provide the most notable example for this in the Swiss case. These elections are organized at the level of cantons, leaving cantonal party sections with a decisive say in the selection of candidates and a commanding role in election campaigns in their constituencies. Parties at lower levels are thus cross-pressured. On the one hand, their programmatic activity is at least partly shaped by the specific regional conflict structure, the particular economic situation of the canton and the constituents' preferences (Moon and Bratberg 2010). On the other hand, however, it is likely that the stronger national arm of a party will attempt to impose its orientation on the subnational parties' preferences. In the case of programmatic differences among the factions of a party, unity is therefore more difficult to maintain (Lindstädt et al. 2011).

It has often been noted that the key to an effective campaign is an affirmative, consistent message, that distinguishes a party from the programmatic efforts of its opponents (Shaw 1999). This is mainly because messages emphasizing specific issues are capable of activating latent predispositions in the electorate (Gelman and King 1993; Holbrook and McClurg 2005). Political parties can thus have a decisive impact on the electoral outcome by shifting the center of attention toward issues on which voters perceive them to be competent (Zaller 1992). A necessary precondition involves the exhibition of a high level of cohesiveness across party communications (Carey 2007; Traber et al. 2014). If the campaign efforts of individual politicians are uncoordinated or only loosely coherent, this signifies to voters that the party is afflicted by internal divisions, and most likely not capable of shaping policy outcomes.

Nevertheless, it still makes sense for some candidates to defect from the party line (Tavits 2009), especially if the party is ideologically heterogeneous and must represent a variety of interests characterizing local voters or interest groups (Bailer 2017). Due to the highly erratic nature of Twitter communication, this micro-blogging service might be regarded as the predestined channel for mavericks trying to win votes by deliberately departing from the party line.

We can therefore formulate two specific expectations for the analysis. On the one hand, we anticipate that the lower the level of government, the more difficult it is for parties to keep messages coherent across their accounts. *In other words, we would be surprised if the national level didn't have more organizational and programmatic coherence than the lower levels of government.* This is because politicians campaigning at the regional or local level are more likely to be influenced by the regional or local conflict structure, such as particular political traditions, a specific economic situation or distinctive voter preferences (Müller 2013; Van Houten 2009). Well-known divisions in Switzerland exist between the language regions on questions related to European integration—with the French-speaking regions, until recently, assuming a pronounced pro-European stance. There are also fierce conflicts in the country's largest party, the Swiss People's Party, between the right-wing populist factions led by the cantonal section of Zurich and the more moderate factions led by the Berne section, which even resulted in the secession of the moderates into a new center-right party. Hence, variations in the structural conditions and preferences across units of a federation can make it reasonable for a party to mount a diverse electoral campaign (Verge and Gómez 2012; Bailer 2017).

The probability that a party will show unity in its Twitter campaigning also depends on demand, i.e. how extensive social media usage for political activities is among the voters of this party (Daniel et al. 2017). The larger the potential audience of a party's communication on Twitter, the greater the chances that such campaigning will have a lasting effect on party attachments (Selb et al. 2009; Stromer-Galley 2014). A party with an already large supporter base on Twitter must therefore care about organizational and programmatic coherence. Specifically, we can formulate our second expectation as follows: *the networks from parties with a large support base on Twitter send more re-tweets from their own ranks and send more similar messages.* In Switzerland, the parties with a large support base on Twitter tend to be smaller (in terms of electoral support), newer and more leftist parties, like the Social Democrats (SPS), Greens (GPS) and Green Liberals (GLP), because their constituencies are typically both younger and more attentive to social media (Wueest and Mueller 2014). If the audience on Twitter is small, relative to the parties' overall voter base, the primary goal of these parties on social media should be to incite and attract voters from different societal groups than their mainstream voters (Cardenal 2011). This is most likely the case for the more traditional, conservative or center-right parties in Switzerland, such as the Liberals (FDP), the Christian Democratic People's Party (CVP), the Swiss People's Party (SVP) and the Conservative Democratic Party (BDP), whose voters share a comparatively low level of interest in social media and are also members of an older age group.

11 Data

Measuring the behavior of the political elite usually involves survey data or roll-call votes (Clinton 2012; Bartels 1991; Bafumi and Herron 2010; Bailer 2017). While conducting elite surveys is costly and conceptually challenging, roll-call data is usually only available for a restricted group of politicians, including members of national parliaments. Other alternative data sources include campaign finance data (Bonica 2014), but in some countries, including Switzerland, this data is not publicly available. We suggest that studying political communication on Twitter represents a viable alternative to tapping into the behavior of the political elite. The collection of data from this micro-blogging service requires comparatively little effort, because it is public to a large extent and can be collected on a large scale via application programming interfaces (API) (see Barberá 2015). Twitter is widely used and offers rich information on political campaigns. Skeptics, however, highlight the potentially high selection bias on Twitter (Pennacchiotti and Popescu 2011) and Jungherr et al. (2012) observe that the only way to achieve an accurate prediction from Twitter data is by accurately identifying a sample that includes the users of interest. This is why we follow a position-based approach in order to systematically trace Twitter users who are relevant to Swiss politics (see Marin and Wellman 2011).

The myriad of Twitter accounts and their highly unstructured descriptions rendered the identification of relevant users the most difficult challenge associated with the data collection. We started out by hand-compiling an initial set of 157 Twitter accounts, which comprised all representatives of the Federal Assembly, all Federal Councilors as well as the official national accounts of the seven most important parties in Switzerland, provided they had a Twitter account in 2015 (see table 25 in the appendix). In four chain-referral extension rounds, using the Rest API of Twitter, we extended this initial set to three steps for each round. First, for all previously identified accounts, the users that follow these accounts as well as the users that are followed by these accounts are collected. Second, a keyword list that contains all names, abbreviations, and paraphrases of Swiss parties, as well as all official employment titles of Swiss politicians in the three official languages of Switzerland (Italian, French, and German), is matched to the Twitter biographies of the roughly 200,000 accounts retrieved in each extension round (see table 26 in the appendix). Finally, all keyword hits are manually checked to confirm their relevance.[11]

Because of its institutional setting, such as the mostly open party lists, and cultural diversity, such as the divide between the conservative countryside and progressive cities, Switzerland has a heterogeneous party landscape. The four parties represented in government, the Swiss People's Party (SVP), Social Democratic Party (SPS), the Liberals (FDP) and the Christian Democratic Party (CVP), also dominate the Council of States (upper house), while the party landscape in the National Council (lower house) is shaped by as many as eleven parties. Given that the network estimates require a statistically significant minimum number of accounts, we restricted our analysis to the eight parties that gained at least two percent of the votes in the election to the National Council. Ultimately, we were able to include 1,341 accounts in the analysis.[12] Subsequently, we manually supplemented the accounts with the following data: canton of residence, party affiliation, gender, institutional level of political activity (national, cantonal and municipal) and political function within the party (party functionaries, elected members of legislative and executive bodies, and party members with no other political function). For this annotation, we relied mainly on official sources, such as election records or protocols of public assemblies. Table 28 in the appendix gives an overview of these indicators.

We also retrieved all tweets sent by these accounts for the period from August 1, 2015, to election day, October 18. The most intense phase of federal election campaigns in Switzer-

11. The information we use for identification is entirely derived from the initial self-declaration of users—the short biography Twitter users can add to their profile—to be aligned to a political party or to hold a political office. However, since we are specifically interested in the campaigning activities of political parties, we think that data collected from Twitter accounts of party members who do not indicate their political affiliation could be even worse than this risk of missing data. These users most likely use their account for purposes other than politics, which possibly leads to non-random biases in terms of campaign activities.

12. Table 27 in the appendix lists the main statistics of the resulting network.

land begins on the 1st of August, the national holiday, on which every politician is obliged to give one or more speeches on the state of the country. The volume of communication varies greatly across the Twitter accounts, but each account had at least one tweet in this period. After filtering for the three national languages, the corpus comprises 129,271 tweets: 95,495 written in German, 30,684 in French and 3,092 in Italian.

12 Measurement Strategy

Building on the following and communication patterns in our sample of party-affiliated Twitter users, we develop three network-based statistics in order to measure the organizational cohesion and programmatic coherence of Swiss parties (see table 3 for an overview). First, we use the follower network to measure organizational cohesion. The follower network shows how well parties are able to connect affiliated accounts. Ideally, a party wants to ensure that each affiliated account follows as many other affiliated accounts as possible, because this ensures that all affiliated accounts stay abreast of each other's status updates. Moreover, the number of followers is often used as a status symbol on Twitter, which is why having affiliated accounts that follow each other is a simple, yet effective, way of boosting parties' Twitter reputations. For this indicator, we directly analyze the network data retrieved from Twitter.

Table 3: Measurements and indicators used in the analysis.

Concept	Statistic
Organizational cohesion	Followers
Programmatic coherence	Re-tweets
	Tweet similarities
Connectivity	Density
Hierarchy	Betweenness centralization

Second, we measure programmatic coherence by computing two indicators from the content of tweets sent by party-affiliated users. For one indicator, re-tweets, we use a mechanism specific to the Twitter platform. Re-tweets allow users to recycle the tweet of another user; to wit, the content of the tweet is copied. This mechanism allows a Twitter user to re-broadcast status updates of other users to their followers, which represents a powerful way of increasing the reach of the original message.[13] As the original message is not altered, the communication is maximally cohesive. For the analysis, for each user, we count which other user(s) they re-tweeted and how often. We standardize these counts by dividing them by the overall number of re-tweets by a user.[14] A connection in the network is then given whenever a user re-tweeted another user at least once. The weight—or strength—of this re-tweet connection is determined by the standardized re-tweet counts.[15]

To compute the second indicator of programmatic coherence, we calculate the similarity between party-affiliated accounts' communication as follows. First, we machine translate the tweets from all three languages (French, German and Italian) into English.[16] Second, we build a weighted bag-of-words representation[17] from the combined tweets of every user.[18] In a final step, we compute the cosine similarity between the word distributions of all users.

We use two statistics from social network analysis to assess the degree of unity across the three indicators just introduced (see table 3). All four indicators have the structure of a social network, although they consist of slightly different types of network data. The first indicator, based on the follower network, is a directed unweighted network, where directed means that the relationship between two network members can—but need not be—reciprocal. The indicator constructed from the re-tweets has an even more obviously directed nature, but the relationships are weighted by the fraction of re-tweets. Finally, the text similarities form

13. There is an important caveat when using re-tweets as proxies for agreement or proximity between Twitter users in the way proposed here. Re-tweeting other users' content does not necessarily indicate agreement with it. However, previous analyses suggest that most re-tweets happen within rather than across party lines (Authors). Because we only use the re-tweets from the same party in the computation, and because intra-party politics in Switzerland does not tend to be very contentious, we can reasonably assume that re-tweets of other party users indicate support or agreement.

14. We also calculated a similar measure for the @-mentions—other Twitter users can be mentioned explicitly in a status update by including an @ and their Twitter user name. Because the results are highly similar compared to the re-tweet network, we only report the analyses using this measure in the appendix.

15. For this standardization, the relative frequency with which this particular user was re-tweeted is used. The weight thereby is a positive number (as a number of 0 results in no connection in our definition above) with a maximum at 1. The maximum weight of 1 occurs if only the messages of a single other user were re-tweeted. In this case, the connection to this users has a weight of 1 and no connection to other users exist.

16. Some research suggests that the Google machine translation service yields satisfactory results (Lucas et al. 2015; De Vries et al. 2018). Our own experience is mixed, which is why we use the recently released *DeepL* Translator instead.

17. In such a representation, a text is represented as the multiset (or bag) for the number of word occurrences. When building this representation, we weighted the number of word occurrences by the document frequency over the inverse document frequency as well as by a topic model in order to make the distribution less skewed.

18. An aggregation of Twitter data is usually recommended for text mining (see Yan et al. 2013).

an undirected and weighted network.

We use the density and the betweenness centralization as the main indicators of two different aspects of within-group unity. As we are interested in intra-party unity, these network statistics are computed on party-specific sub-networks. To wit, we look at each party individually, only considering the connections between its affiliated accounts and ignoring all connections to accounts of other parties.

The density of a network is the fraction of possible connections that are actually present (e.g. Wasserman and Faust 1997). One important caveat to keep in mind when comparing the densities of *social* networks of disparate sizes is that, all else being equal, smaller-sized networks tend to have higher densities than larger-sized networks (Scott 2017). In our interpretation of the results, we will therefore pay attention to comparing only the densities of parties with similarly-sized networks.[19]

The second statistic we compute is the betweenness centralization coefficient. Relative to other definitions of centrality (see Freeman 1979), betweenness centrality is conceptually closest to the type of within-group unity we discuss above, because the betweenness centralization coefficient can be understood as a measure of hierarchy, i.e. how strongly direct connections in the network depend on a small set of actors (Freeman 1977, 39). The betweenness centralization coefficient is computed from the actor-level betweenness centrality, by averaging the differences between the most central actor and every other actor in the network. This number is then standardized to a range of one unit, such that it is 0 when every actor in the network has the same betweenness centrality and 1 when the only connections in the network are between a single actor and every other network member.

We obtain uncertainty estimates for the two network statistics from the non-parametric bootstrap. Resampling a subset of the network under study demonstrates the sensitivity of our results to the network boundaries (Galaskiewicz 1991; Costenbader and Valente 2003). Although we compute network statistics for party-specific sub-networks, the identification of the network sample happens at the level of the full network. Therefore, resampling is also

19. Computing the density of weighted networks can be difficult because the standardizing factor depends on the scale of the weights (Wasserman and Faust 1997, 143). However, because all our weights take on values between 0 and 1, we can compute the upper bound for the sum of weights. More precisely, we can use the standard upper bound for the number of connections in an unweighted network of n users which is $(n \cdot (n-1))/2$ or $(n \cdot (n-1))$ for undirected and directed networks, respectively.

done for the whole network. For the main analysis, we report uncertainty estimates from 1,000 repetitions of resampling 95% of the network.[20]

13 Results and Discussion

13.1 Network Structure

The empirical analysis starts with an initial overview of the network structure of the Swiss parties' Twitter campaign in the run-up to the Federal Election of 2015. Figure 4 presents a visualization of the political Twitter network in play for the 2015 Swiss federal election campaign. The general arrangement of the nodes and communities is based on two nested Fruchtermann-Reingold layouts–one applied to the communities, and a second applied to the nodes within their community. The communities were detected via the cluster algorithm for large networks proposed by Clauset et al. (2004). An overview of the most important characteristics[21] of the communities is provided in table 4.

The size of the vertices in figure 4 represents the betweenness centrality of the party accounts. Generally, the inequality in this centrality reflects the rather hierarchical nature of the Swiss party network. Although roughly half of the accounts are highly central—typically party presidents, Federal Councilors, National Councilors, Councilors of State or accounts of the national party offices—there are also many peripheral accounts with only sparse connectivity to the network.

The four communities reflect the two most salient divides characterizing the Swiss political Twittersphere. On the one hand, there is a clear ideological left-right divide within the network (see table 4). Though the first two communities are clearly shaped by politicians from leftist parties—the Social Democrats (SPS) and Greens (GPS), community 3 is almost exclusively occupied by centrist parties–the Christian Conservatives (CVP) and Green Liberals (GLP). Community 4, finally, mirrors the right pole of the spectrum, as the overwhelming majority of politicians from the Liberals (FDP) and the right-wing populist Swiss People's

20. We also ran all estimations where we never dropped the initial set of accounts that started our chain-referral sampling from the bootstrap samples, because we wanted to test what happens when we mirror the data collection process. The interpretations of the results are largely unaffected by this change.

21. Only tabulations are shown for the main indicators whose relationship with the community memberships is significant in a χ^2-test.

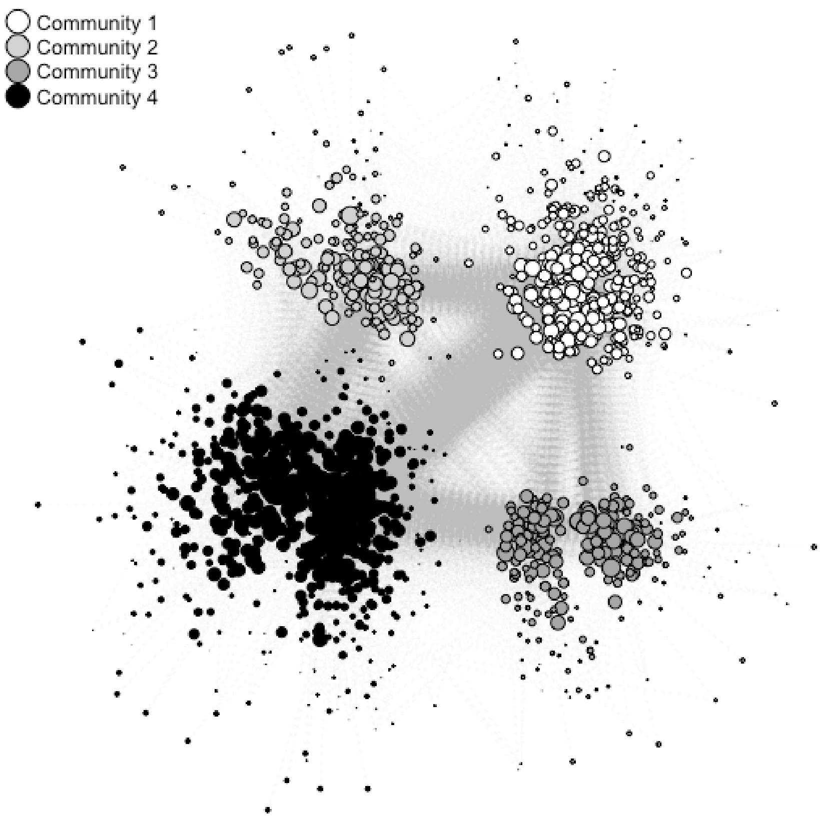


Figure 4: The political Twitter network in the 2015 Swiss federal election campaign. The nodes indicate the Twitter accounts and the edges show their follower-relationships. The size of the nodes signifies an account's betweenness centrality.

Table 4: Relevant characteristics of the communities shown in figure 4. Frequencies in %.

Party[a]	BDP	CVP	FDP	GLP	GPS	SPS	SVP
Community 1	2.8	4.0	1.6	5.7	65.8	69.3	1.5
Community 2	6.9	16.6	2.0	4.1	32.3	29.1	1.9
Community 3	0	47.1	1.6	88.5	0.6	0.3	3.4
Community 4	90.3	32.3	94.7	1.6	1.3	1.4	93.2
Language[b]	DE	EN	FR	IT			
Community 1	38.9	29.3	5.5	16.3			
Community 2	1.0	8.6	45.5	4.7			
Community 3	23.9	19.8	1.9	7.0			
Community 4	36.3	42.2	47.1	72.1			

Note: Only tabulations are shown for the main indicators whose relationship with the community memberships is significant in a χ^2-test.
[a] BDP = Conservative Democratic Party; CVP = Christian Democratic People's Party; FDP = the Liberals; GLP = Green Liberal Party; GPS = Green Party; SPS = Social Democratic Party of Switzerland; SVP = Swiss People's Party.
[b] DE = German; EN = English; FR = French; IT = Italian.

Party (SVP) are found in this community.

A second fundamental, and quite particular, characteristic of political communication in Switzerland is also reflected in the community structure displayed in figure 4. The Swiss media system is generally separated according to language borders, which is manifested by different private media outlets, and also expressed by separate public broadcasters for every language region. The social media platform Twitter is no exception to this (see table 4). Although Twitter accounts are not as strongly separated by language as by ideology, the decision remains significant. All other indicators for the accounts, such as gender or political function, do not significantly induce differences among the communities.

The language-related fragmentation in the Twittersphere even operates within single parties. The French-speaking accounts from most parties are, accordingly, mainly grouped into community 2, while their German- and Italian-speaking party colleagues have a stronger presence in communities 1, 2 and 3.

13.2 Organizational Cohesion

The analysis proceeds with an examination of the follower network in figure 5.B, which reveals several interesting patterns for the size of the Twitter network, its density and betweenness centralization.[22] First, it seems obvious that the cantonal and local follower networks are bigger than the national networks, simply reflecting the upwardly-narrowing hierarchy of the party organizations. The local networks, however, are not always larger than the cantonal ones. Cantonal parties are as important, sometimes even more important than the local ones, for the BDP, GLP, GPS, SVP and CVP. This result may be partly due to our sampling strategy, which started with a federal-level sample of users. However, it might alternatively reflect the importance of the cantonal level for Swiss party politics.

Network density provides information on the horizontal component of organizational cohesion—or in other words, on the average connectivity in the network. The results for the densities of the follower networks provide strong supportive evidence for our first expectation that it is tougher or less desirable for parties to control the campaigning of lower-level politicians. Although the network densities of the cantonal and local networks are close to the overall density across all parties, the density of the national network is clearly higher for all parties except the SVP. Hence, the national networks in the Swiss political Twittersphere tend to be much more tightly connected in organizational terms. By contrast, there seems to be more room for independent network building at the lower levels of Swiss politics.

As briefly discussed in the section on measurement strategy, it is always necessary to acknowledge the positive correlation between the density and the size of the network if the network densities are compared across parties.[23] With this in mind, the density of the SVP user network should be much higher, relative to its respectable size. In terms of the follower relationships among its accounts, the SVP is evidently only loosely organized. The clearest outlier at the other end, however, is the national network of the SPS. With a Twitter network

22. We only provide evidence for the intensity of follower relations in this paragraph, but it becomes apparent that the Twittersphere in Switzerland is quite representative of the electoral landscape, with one important exception. Besides the SVP, arguably the most important Swiss party, the ranking of all parties corresponds to common measures of party strength, such as the number of seats in the National Council (see also table 28 in the appendix).

23. Much of this correlation is purely data-driven and not particularly interesting, substantively (Scott 2017). Therefore, it does only make sense to compare the density of smaller parties, such as the BDP, with the density of the larger parties if the size of the network is considered at the same time.

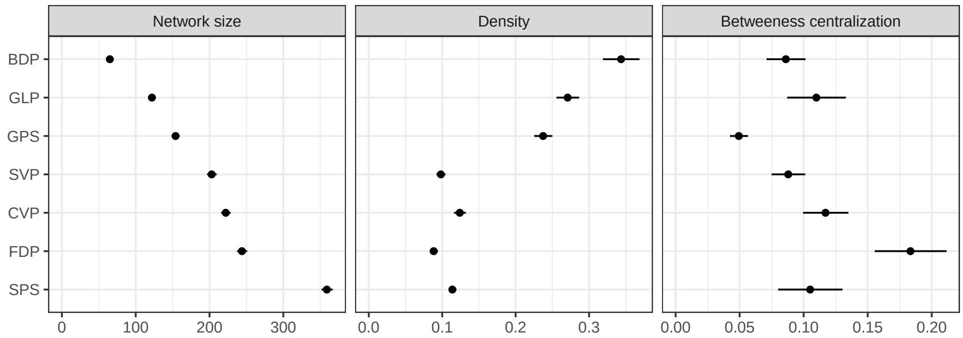


Figure 5.A: Organizational cohesion over all levels.

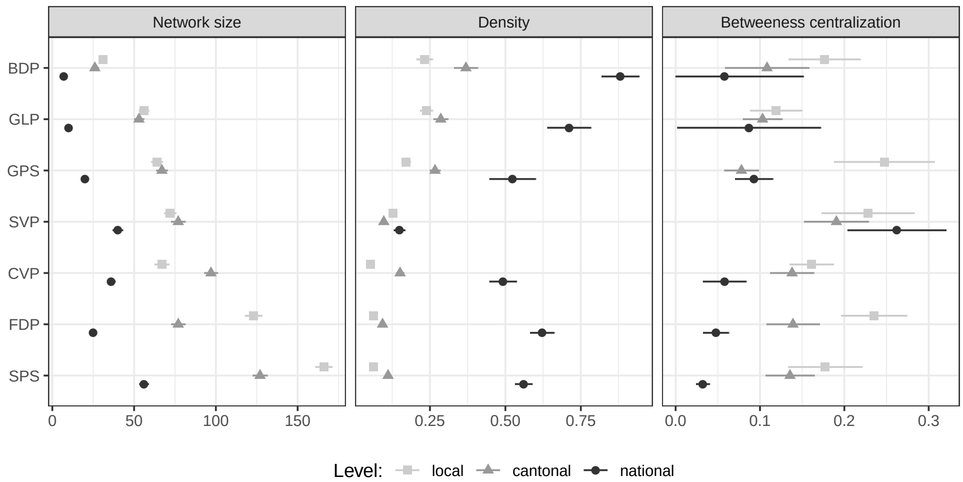


Figure 5.B: Organizational cohesion by level.

that is very large, the Social Democrats are able to maintain a level of connectivity that rivals parties with much smaller-sized networks. This is a clear indication that the SPS is more highly organized than the other parties.

The second network measure, betweenness centralization, indicates the degree to which a party is centered around a few highly connected users. It is therefore suited to uncovering the hierarchical component of organizational cohesion. As the results in figure 5.A show, the larger parties in terms of their follower networks are generally more hierarchically structured than the smaller parties, especially the GPS.

Two further patterns are noteworthy. First, at the national level, the networks of the three Federal Council parties, CVP, FDP and SPS, are among the least hierarchical. This probably reflects the consensus-oriented nature of federal politics in Switzerland, with pragmatically, rather than ideologically driven policy coalitions. Similarly interesting is the result for the national SVP, the only other party that regularly elects representatives to the Federal Council. It is a distant outlier from the general trend, which seems to confirm its exceptional position in the political system of Switzerland. The second interesting pattern is that the two parties that are farthest to the right in the political spectrum, the FDP and SVP, in general are more hierarchically structured than the other parties.

13.3 Programmatic Coherence

The results for measuring programmatic coherence via re-tweets and tweet similarities are displayed in figures 6.A and 6.B. As for the densities of the re-tweet networks, the parties can be split into one large group with broadly similar network sizes, accompanied by two outliers. The BDP is the downward outlier with the smallest network. The density of its re-tweet network, however, is not significantly larger than the density of the re-tweet networks of the other parties. This indicates that Twitter users from the BDP use re-tweets comparatively less to refer to the content of their party-affiliated accounts.

The large group of medium-sized parties includes the GPS, GLP, SVP, CVP, and FDP. The network sizes of all of these parties are quite similar, despite smaller differences. The densities of their re-tweet networks are also similar, with the minor exception of the CVP. Relative to the size of their network, the users of this centrist-conservative party are only

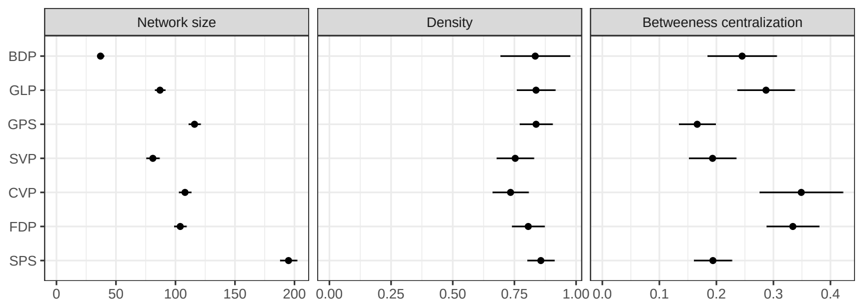


Figure 6.A: Programmatic coherence: Re-tweets over all levels.

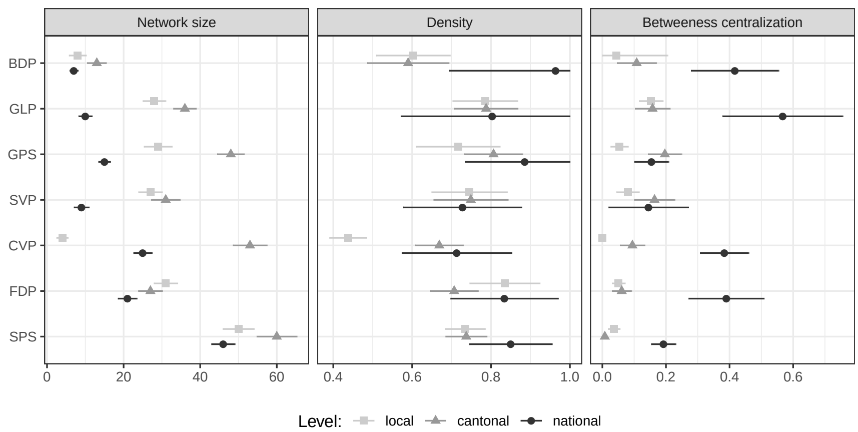


Figure 6.B: Programmatic coherence: Re-tweets by level.

weakly connected in horizontal-organizational terms, not only (but most notably) at the local level. The upward outlier is the SPS, which has about twice as many users in its re-tweet networks as the parties in the large middle group. With associated densities only slightly smaller than the counterparts from this comparison group, the SPS, despite its larger-sized network, seems to be much more horizontally organized relative to those five parties.

Turning to the hierarchical structure of the re-tweet indicator, a general difference among the betweenness centralization in the follower networks is that it tends to be more weakly connected at the local level. In five of seven cases the difference is significant and the national level displays higher values in terms of betweenness centralization. In the analysis on the follower networks, the local level users, if they exist, stand out as more hierarchically structured, and higher-level users have more organizational cohesion. In terms of programmatic coherence, the accumulation of centrality by a few users is more pronounced at higher political levels.

Moreover, there are basically two groups of parties exhibiting broadly comparable levels of hierarchy in their re-tweet networks. The first group includes the BDP and GLP, two centrist parties with a smaller network size, as well as the CVP, FDP and SPS. The re-tweet networks of all these parties are characterized by a strong hierarchy at the national level and much less programmatic coherence at the cantonal and local levels. Local and cantonal users from these parties are thus substantially freer in their re-tweet behavior than their national counterparts, who seem to coordinate their campaigns with much more effort.

The other group of parties, the GPS and SVP, have neither high centralization scores generally nor large variation across the political levels. This indicates that there are no Twitter users in these parties setting the pace for the other party-affiliated accounts. Because there is evidently less pressure to spread the party message coherently, there is potential for single users to strategically deviate from the party line. This is somewhat surprising, as far as the SVP is concerned, since users from this party have shown exceptionally strong organizational cohesion.

Figures 7.A and 7.B display the results from the second indicator of programmatic coherence, based on the text similarity of the tweets sent by party-affiliated Twitter users. In general, the national level of party politics is much more connected in terms of tweet simil-

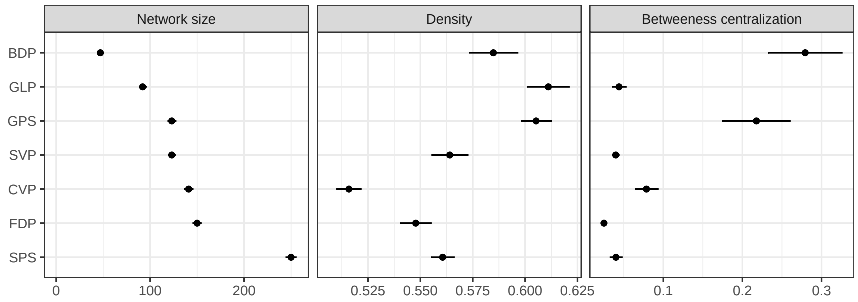


Figure 7.A: Programmatic coherence: Tweet similarity over all levels.

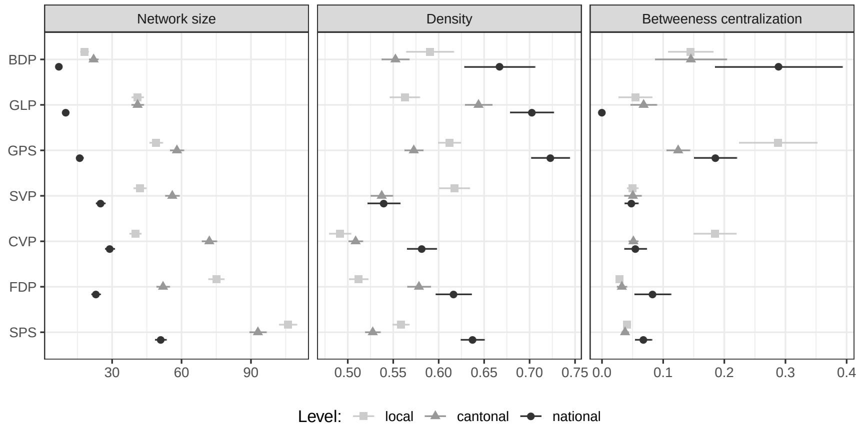


Figure 7.B: Programmatic coherence: Tweet similarity over all parties by level.

arities. As with the betweenness centrality of the re-tweet networks, coherence among the cantonal and local politicians is considerably weaker. This is a clear indication that the national campaign message is adapted to regional peculiarities in all parties.

It is worth noting the lack of an implicit relationship between the density and the size of the network in this case. However, for consistency, we do include the network size in the graph. The reason for the independence between the two statistics is that the density, in this analysis, simply reflects the average distance between two users from a given party.[24]

With the ability to compare across all parties in the analysis, the smaller parties—BDP, GPS and GLP—are revealed as having a comparatively high degree of text-similarity-based coherence. The bigger parties—SVP, CVP, FDP and SPS— tend to have a lower degree of programmatic coherence in this regard. This difference between the two mentioned groups

24. This follows from the way we compute the density, which sums up the weights, which are the text similarities as described in the section *Measurement strategy*, and divides by the number of potential connections, which equals the number of terms in the sum.

of parties is especially pronounced at the national level, although national users from the SPS exhibit similarly high levels of consistency, relative to the national users from the GPS, GLP, BDP. For the other larger parties, most notably the CVP, there is more variation in the campaign message. A possible interpretation is that these parties must cater to a more diverse constituency during the election campaign. Hence, Twitter seems to be a feasible campaign platform for addressing this increased diversity.

Two of the smaller parties, GPS and BDP, tend to have a slightly more hierarchical structure in their text similarity networks, as the final graph on the betweenness centrality shows. The other parties have similarly low levels of centralization. Although this pattern does not hold for the GLP, it nevertheless seems to be the case that the communication of smaller parties across all political levels is more heavily centered on a few users. The Twitter communication of bigger parties, by contrast, is more evenly distributed across the users.

14 Conclusion

Existing studies on party unity in Switzerland have almost exclusively focused on representatives at the federal level (see, e.g. Traber et al. 2014). We argue that these findings can only grasp part of the story, since party unity in Switzerland is constantly under strain by regional divisions across the different levels of government. By using data on Twitter accounts, we can extend considerably the number and types of politicians considered in the analysis, and thus draw conclusions beyond the narrow national realm. Such a research strategy therefore seems generally promising for studying party unity in other strongly federalist states such as Germany or Belgium, where subnational party factions are comparatively powerful. Hence, we can expect similar patterns across the electoral campaigns in such comparable countries.

In a first step, our descriptive analysis largely confirmed common knowledge about 'offline' politics in Switzerland. On the one hand, we found a clear congruence between the parties' number of Twitter accounts and their strength in the national parliament. On the other hand, the general ideological left-right divide of the Swiss political system, as well as the usually clear language-specific separation of the Swiss media system, have become evident in how the Twitter accounts grouped into network communities. These findings are also consistent with qualitative evidence stemming from interviews with election campaign managers in the run-up to the 2015 national elections (Fichter and Kohler 2015).

In a second step, we showed that our analysis is, at least to some extent, able to disentangle organizational and programmatic coherence. This is something that previous studies on party unity that focus only on single indicators, such as roll-call votes, cannot achieve. More precisely, we use the follower network as an indicator of organizational cohesion and the re-tweet and text similarity networks as indicators of programmatic coherence. Computing the density and the betweenness centralization allows us to quantify and compare programmatic and organizational unity across parties and levels of government, while the non-parametric bootstrap allows us to incorporate uncertainty to these measurements.

We find that, with respect to most statistics on organizational and programmatic unity, the national levels of the party networks tend to be more tightly connected. Hence, it is clearly tougher for parties to maintain unity in their networks at lower levels. Because cantonal and local peculiarities are often important factors in the political campaigns of lower-level politicians, it is probably even desirable for national party offices to allow them more room to maneuver. Local and cantonal users from these parties are thus substantially freer in their following, re-tweeting and general text messaging than their national counterparts, who seem to coordinate their campaigns with much greater effort. This evidence is almost certainly also due to the nature of the social media platform Twitter. With the brevity of its messages and high-speed dissemination, Twitter promotes a very heterogeneous communication. It might therefore be deliberately chosen by parties in order to leave room for dissenters, who, in turn, might be able to address voters outside their traditional constituencies.

Our comparisons across parties revealed that larger parties are generally less hierarchically structured, in organizational terms, than the smaller parties. This does also hold for the text similarity networks, where the accounts of smaller parties are significantly less equally distributed. One important exception is the national level of the Swiss People's party. This seems to confirm its exceptional position as a highly populist and disruptive force in the political system of Switzerland. Among others, this party is well-known for concentrating on a few media-savvy figureheads.

Generally, however, the Social Democrats are responsible for the most remarkable results. Despite their very large network, the Social Democrats are able to maintain a very high level of connectivity, in terms of follower relations and re-tweets. The SPS is therefore much more organizationally and programmatically cohesive than the other parties. In these times, when Social Democratic parties are struggling in the context of election campaigns throughout Western Europe, this skillful use of Twitter by the SPS may provide a glimmer of hope.

Switzerland is not a particular front-runner with respect to the digitalization of its election campaigns, but also here, all parties except the CVP claimed to have spent considerable financial resources on Twitter during the 2015 national election campaign (Fichter and Kohler 2015). Moreover, each party offered training sessions and workshops on how to use social media. Since then, more politicians have joined Twitter and communicating on social media has become more important. Based on our experience from Switzerland, we can therefore expect social media to become more influential as a complement to more classical campaign strategies also in other established democracies. This offers a wide variety of opportunities for us to extend our research into organizational cohesion and programmatic coherence in further countries and other social media.

Chapter V

From Party- to Candidate-Centred Campaigns

The focus of the analysis in the previous chapter is the party because it plays a decisive role in organising candidates in multi-party election systems (see section 9). The party's goal is to ensure that "the message" is clearly communicated to potential voters (Freelon 2017). As such, the party should be concerned with keeping their candidates "on message". Individual candidates might favour deviating from the party line because, depending on the electoral system, they want to maximise votes for the party but also for themselves. Focusing on candidates' incentive to integrate social media into their campaign—and how they decide to do it—is a natural extension of the party-centred view.

The two following chapters contain an analysis of social media use during the 2019 European Parliament elections. The first step is identifying the social media accounts of candidates across EU countries, which I describe in Chapter VI. This paper discusses how I gathered ballot papers in all EU countries, searched for social media accounts, and ensured the quality of the process. It also gives an overview of the resulting dataset and outlines potential applications. This dataset is the foundation for Chapter VII, in which I analyse how electoral systems incentivise candidates' social media use. Based on theoretical work by Carey and Shugart (1995) and Shugart (2005), I develop and test hypotheses which extend recent empirical work by Obholzer and Daniel (2016) and Daniel and Obholzer (2020).

Chapter VI

(Paper 3) A Comprehensive Dataset of Social Media Adoption in the 2019 European Parliament Elections

Abstract

European Parliament elections are a fruitful ground for the study of candidates' adoption of social media as a campaign strategy. However, taking full advantage of this research opportunity requires high-quality datasets. This paper presents a comprehensive data source covering candidates in all 28 European Union (EU) countries during the 2019 elections. It documents the process of compiling the dataset, starting with the collection of the names of all 15,540 candidates from the 560 parties taking part in the election. Parties and candidates within these parties were then included in a search for social media accounts, leading to the identification of 2,368 Twitter and 2,626 Facebook accounts belonging to 3,862 candidates in the search. To show the range of potential applications of this dataset, I present descriptive and model-based evidence of social media adoption during the 2019 European Parliament elections, and I outline additional research areas in which this dataset could be useful.

15 Introduction

European Parliament (EP) elections represent a unique opportunity for comparative social media research because they are held under common rules and share common issues of contestation at the level of the European Union (EU), while still being conducted separately by each EU country with its own country-specific election context. Such research requires appropriate and comprehensive data sources to use the full power of this opportunity by comparing all EU member states. However, datasets which enable this type of research are difficult to collect due to the extent of the European Parliament elections. The work of Nulty et al. (2016) shows how compiling a comprehensive dataset for the 2014 European Parliament elections lead to important comparative findings about the adoption and use of social media platforms. This paper describes the identification and collection of social media accounts for the 2019 European Parliament elections.

16 Creation of the Dataset

The creation of the dataset proceeded in three stages. Firstly, information on every candidate across all EU countries needed to be collated. Secondly, candidates' social media accounts had to be identified. Finally, the dataset was enriched with data from external sources to increase its usefulness for different types of analysis.

16.1 Collecting the Names of Candidates

The main challenge when creating a list of every candidate in a European Parliament election is that each member state runs separate elections. The EU only sets the time and requires them to have some form of proportional representation. This is an issue for research because member states have different rules for the registration of parties and the publication of lists. Finding candidates' social media accounts first and foremost requires a list of all candidates standing for election. The problem here is that official ballots are only posted shortly before the election. The EU requires lists to be published at least four weeks before election day, which leaves only a short window for the collection of social media data. Table 29 in the appendix shows the sources and collection date for each country.

Identifying the correct official websites which publish the names of parties and candidates was often time-consuming. Obtaining complete lists of parties and candidates from these websites was possible with standard web scraping tools. For example in Poland, complete candidate lists in all 13 districts were published online by the Polish National Electoral Commission. Lists for each district only had to be retrieved, extracted from the raw HTML, and processed into a standard table. I was able to generate candidate lists in many countries in this way, although additional steps were sometimes required when a particular website did not lend itself easily to web scraping.

Unfortunately, it was not always possible to use lists from official sources because they were published shortly before election day, sometimes by regional election offices, often for ad-hoc districts specifically created for the European Parliament elections, or they were not available in a machine-readable format. Italy is an example of the first two issues. The country is split into five districts for European elections. The starting point for the collection was party websites, which published the names of candidates. However, not all parties had a dedicated website for the EP elections. The next source for finding candidates was national newspapers because they were the only ones that included the names of candidates. While this allowed me to start the collection, these lists needed to be verified via official publications. In this case, the official lists first became available as scans of judicial decisions accepting the candidates/lists for each district separately. Finally, an official table was released, which I could use to verify the initial lists. Another example of these problems is the case of the United Kingdom. The election was run entirely from regional election councils for the 12 districts specifically created for the EP elections. There was no official national-level source, and the councils all had separate websites where they published scanned versions of the official ballot paper. The complete list of candidates was compiled manually from these official ballot papers.

The overall number of parties[25] and candidates across 28 countries is extremely high. Table 5 shows that 15,540 candidates from 560 parties competed in the 2019 election. The number of parties ranges from 7 in Austria to 41 in Germany, and the number of candidates

25. I refer to them as "parties" although a number of electoral lists comprise candidates from different parties.

from 41 in Malta to 2,686 in France, where every party has to supply a full list of 79 candidates.

Table 5: Overview of lists and candidates selected for collection.

	All		Selection	
Country	Parties	Candidates	Parties	Candidates
Austria	7	260	7	124
Belgium	21	153	12	118
Bulgaria	27	318	7	74
Croatia	33	396	6	40
Cyprus	16	–	5	30
Czechia[a]	39	841	8	158
Denmark	10	135	7	100
Estonia	10	66	5	32
Finland	18	269	7	70
France	34	2686	8	343
Germany	41	1380	7	250
Greece[a]	40	1195	6	233
Hungary	9	291	5	54
Ireland[b]	–	59	–	59
Italy[c]	18	1076	4	295
Latvia	16	246	9	54
Lithuania[a]	16	301	8	60
Luxembourg	10	60	6	24
Malta[b]	–	41	–	24
Netherlands	16	392	10	86
Poland[a]	9	866	9	458
Portugal[a]	17	483	7	75
Romania[a]	23	695	7	137
Slovakia	31	349	10	70
Slovenia	14	103	7	26
Spain[a]	32	1907	9	180
Sweden	31	518	8	144
United Kingdom	14	544	14	544
Total	560	15540	198	3862

[a] Candidate count includes substitute/additional candidates.
[b] Country uses Single Transferable Vote (STV).
[c] Candidates can stand in more than one electoral district.

As in many multi-party proportional representation elections, there are many parties and candidates whose chance of being elected is very low. This includes special issue parties such as "The Violets" (Die Violetten) in Germany or "The Forgotten of Europe—Artisans, Traders, Liberal Professions and Independents" party (Les oubliés de l'Europe—Artisans, commerçants, professions libérales et indépendants) in France. As described above, the large number of parties in the election requires the selection of parties with at least some chance of gaining a significant number of votes. For this collection, I selected parties based on pre-election polls. The selection rule was that a party needed to have a projected vote share equal to at least one seat per country. Even in parties likely to get at least one seat, many candidates do not have a realistic chance of being elected. They rank so far down on the list that the party would need to win most or all of the seats for them to be elected. To effectively use available resources, I also selected candidates within lists. In line with using predicted vote and seat share from national polls, I selected twice as many candidates as a party's projected seat share.

16.2 Collecting Social Media Accounts

Procedure

There are two challenges when searching for candidates' social media accounts. First, it is crucial to find social media accounts even if they are not easily found on, for example, party websites. Second, the social media accounts have to be used for campaigning in one way or another. The implicit assumption in studies of the adoption of social media is that some party or campaign-related activity is happening on these accounts. While analyses of the *content* of social media posts usually use a classification method to filter out posts without programmatic messages, analyses of who has a social media account usually retain all the accounts they find.

With these two considerations in mind, the search for accounts on Twitter and Facebook[26] proceeded in the following way. First, party websites were searched for links to social media accounts because a few parties had dedicated websites for their EP candidates, which could easily be copied. In most cases, no party lists were available, so I checked Wikipedia entries for links to social media accounts. Naturally, only the better known candidates had such an entry, but it helped to make quick progress for those candidates. If the candidate did not have a dedicated Wikipedia entry, I continued by searching Google with the name of the candidate and the short form of the party name. For example, to find the social me-

26. The data also contains Instagram accounts for certain candidates, which were not collected systematically, however.

dia account of Sergio Coronado, the 14th candidate on the list of the “Unbowed France” (“La France Insoumise” in French) party in France, I constructed the following query URL:

`https://www.google.fr/search?pws=0&cr=FR&q=Sergio+Coronado+Insoumise`

The query `q=Sergio+Coronado+Insoumise` is composed of the candidate’s name and the party’s short name, in this case “Insoumise”. The other parts of the URL are to instruct Google to use non-personal searches (`pws=0`), that is, not to use the history of previous queries to tailor the results to the user, and to bias the results towards France as the website’s source (`cr=FR`).[27] This led to results which often included the (main) social media account, as in this example, which returned the profile on the party website and the Twitter account among the first few results. If the results did not include social media accounts, I added the specific social media platform as part of the query, for example, “Sergio Coronado Insoumise Twitter” and “Sergio Coronado Insoumise Facebook”. If there was still no result, I used the search fields on Twitter and Facebook directly with the name of the candidate and the party.[28] [29]

As the search proceeded, especially the searches directly on Twitter/Facebook, it became harder and harder to clearly identify whether an account actually belonged to a candidate with the given name, especially because some names are relatively common. If there is no indication of a party affiliation anywhere in the description, it becomes hard to attribute the accounts. In such cases, I looked at the posts and sometimes even pictures they had posted to find an indication of the relevant party. However, in many cases, this was unclear. I decided not to affiliate an account with a candidate unless there was a clear indication, such as a function/party name somewhere, that this account belonged to the candidate and was even used in a programmatic function.

27. It is important to note that the “country bias” functions as a hint for Google Search and does not restrict the results to this country only. Nevertheless, it generally worked well for generating successful searches in this case.

28. The specific URLs are https://twitter.com/search?f=users&q=Sergio+Coronado+Insoumise and https://www.facebook.com/search/top/?q=Sergio+Coronado+Insoumise.

29. Candidate and party names in Greece and Cyprus, which use Greek script, and Bulgaria, which uses Cyrillic script, had to be transliterated to the Latin script to generate these search queries. I used standard systems for romanizing and latinizing these names in both cases. I employed two research assistants who are native speakers of the respective language to run these queries and select the results.

The collection process was implemented in the way described in section 16.2. I started collecting lists of candidates in April 2019, beginning with the most populated countries that had some data available at the time, such as France and Germany. However, due to the rules of European Parliament elections—electoral lists only have to be officially published four weeks in advance—this only worked for a few countries, and even then, I had to resort to party websites or newspaper articles. Instead of continuing to work with different unofficial sources, I decided to prepare country-level data for the collection of individual social media accounts based on the few lists I was able to retrieve. Collating the lists resumed towards the end of April and was mostly finished in early May. Simultaneously, three research assistants and I started to go through the candidate lists and used the online searches I had set up to look for social media accounts.

The protocol of checking twice as many candidates as the projected vote share of a party was followed in most countries. However, there were a few countries where the social media accounts of all candidates from a selected party were collected. These exceptions include Ireland and Malta, which use a single transferable vote system, and the selection rule for candidates based on projected party vote shares is impossible. Four additional countries were fully checked because time allowed it. Finally, in the United Kingdom, which had not planned to run the election because it had expected to leave the EU beforehand, and where new parties and ad-hoc electoral districts made it impossible to select candidates based on projected vote share, the social media accounts of all candidates were collected.

16.3 Adding External Data

Additional data was added after the collection of social media accounts was finished. While the ballot paper notes the gender of a candidate in certain countries, it could not be collected from election lists for most candidates. Additional gender information was added through the Gender API online service.[30] This service allows the inclusion of a *country* parameter to give a hint about given names, which might differ in the likelihood of being male or fe-

30. `https://gender-api.com`

male based on the language and country. The names of all candidates were submitted to the service, which returns the classified gender and the likelihood the model assigned to its decision.[31] There were some candidates whose names were classified with a very low predicted accuracy. These are added to the data as "n/a".

Party-related information was added from the ParlGov project (Döring et al. 2020), whose identifiers are used to identify parties. For election lists comprising several parties, the lists were attributed to the main party or the party with the highest national vote share. Almost all parties in the dataset have a corresponding ParlGov identifier, apart from small parties and new parties, which are not (yet) covered. This identifier was then used to add information about party positions—for example, general left–right and support for EU integration—and party performance in national elections—for example, the vote share and whether a party is in the cabinet. At the same time, the data on party positions, which in ParlGov is based on the 2010 version of the Chapel Hill Expert Survey, was updated with the latest data from 2019 (Jolly et al. 2022).

Incumbents, as well as candidates elected in the 2019 elections, were matched with their unique identifier on the website of the European Parliament.[32] It was subsequently used to scrape the EuroParl website and retrieve the candidates' history of service in the European Parliament. Although not done in this case, these identifiers also allow the merging of additional data, such as the committee history or legislative activity.

Finally, the dataset also contains country-level data. Firstly, there is information about social media and internet usage from the Eurobarometer survey (European Commission, Brussels 2019). Specifically, questions QE3_5 and QE3_6 in the survey ask participants how often they use the internet and online social networks, respectively, on a 6-point scale from "Everyday / Almost everyday" to "Never". Two variables are created from each variable indicating the share of respondents per country who answered the highest possible usage, "Everyday / Almost everyday", and the two highest categories, "Two or three times a week" or more. Figures 22 and 23 in the appendix compare the two variables in the Eurobarometer survey

31. Checking the classifier's output against candidates for whom the "true" gender is available shows that the Gender API model is highly accurate.

32. For example, the identifier of Manfred Weber from Germany is 28229. It uniquely identifies him on the website: https://www.europarl.europa.eu/meps/en/28229/MANFRED_WEBER/home.

and the four resulting variables in the dataset. Secondly, the data also contains election outcomes at the country level. Specifically, the number of eligible voters and how many valid votes were cast.

16.4 Validation

To guard against potentially missing candidates by selecting them based on their position on the ballot paper, I randomly sampled 2–3 parties in each country and collected information on all the candidates on the list. Random sampling ensures that a range of different parties are represented in the validation sample, at the expense of potentially sampling very small parties. The validation sample also includes countries where every candidate on every list was checked. This includes countries using STV, such as Ireland, countries with ad-hoc constituencies, such as the UK, and a few additional countries in which RAs had additional capacity, such as Greece. Table 30 in the Appendix shows an overview of how many parties were fully collected.

This data can be used to validate the collection procedure of the dataset itself. Specifically, I can look at the completely collected parties and compare how many candidates have accounts among those candidates who would have been collected in any case and the candidates which were only checked due to the complete collection. The result gives an indication of how many accounts I would have found if I had checked every candidate for every party. The Social Democrats in Sweden were, for example, randomly selected for complete checks. Opinion polls predicted that the Social Democrats would win five out of 21 seats. Checking twice as many candidates would have found that 100% have a Facebook and 90% have a Twitter account. Among the 11 candidates further down the list, only 18% have a Facebook and 9% have a Twitter account. In the case of this party, the validation data shows that almost all accounts were collected following the procedure and only a few accounts would have been missed.

17 Applications

To show the usefulness of the dataset for studying substantive research questions, I show descriptive evidence about the adoption of social media, and I re-estimate a model about the adoption of Twitter during the 2014 European Parliament elections. Finally, I describe additional research areas in which this dataset could be useful.

17.1 Descriptive Overview

Table 6 provides a descriptive overview of the collected data. The first two columns show how many parties and candidates were selected for collection. The next two columns show how many candidates have Twitter or Facebook accounts, respectively. Of the 3,862 candidates who were selected for collection, 2,368 have a Twitter account, and 2,626 have a Facebook page. This overall comparison obscures the fact that there is considerable variation between countries. On one side are countries such as the United Kingdom, France, Spain and the Netherlands, where candidates more often have a Twitter account than a Facebook page. The United Kingdom is a particularly extreme case with more than 80% of the candidates using Twitter compared to only 16% adopting Facebook. On the other side of the spectrum are countries such as Czechia, Romania and Bulgaria, where Facebook is almost universally adopted, while fewer candidates use Twitter. Most countries fall between these two extreme groups with a pattern of moderate adoption rates for both social networking sites but almost always more Facebook than Twitter accounts. Looking only at the extreme examples might suggest that the use of Twitter over Facebook, and vice versa, stems from a divide between Western and Eastern European countries. However, this is only a superficial pattern. Some large countries in Western Europe, for example, Germany or Italy, have candidates who prefer Facebook over Twitter.

A descriptive analysis could, for example, explore the consistency of one of the results in chapter IV which shows that different party families adopt Twitter differently in Switzerland. Extending this analysis to the 2019 European Parliament elections and including Facebook accounts in Figure 8 does not show a clear pattern of adoption of one social media platform over the other. Only the candidates of green/ecologist parties have more Twitter than Face-

Table 6: Overview of collected data.

			Accounts	
Country	Parties	Candidates	Twitter	Facebook
Austria	7	124	43	86
Belgium	12	118	76	89
Bulgaria	7	74	27	61
Croatia	6	40	21	38
Cyprus	5	30	13	20
Czechia	8	158	62	129
Denmark	7	100	63	85
Estonia	5	32	20	29
Finland	7	70	55	61
France	8	343	281	218
Germany	7	250	143	180
Greece	6	233	95	154
Hungary	5	54	21	41
Ireland	14	59	45	38
Italy	4	295	188	255
Latvia	9	54	43	54
Lithuania	8	60	23	63
Luxembourg	6	24	15	23
Malta	2	24	20	23
Netherlands	10	86	68	42
Poland	9	458	261	378
Portugal	7	75	30	50
Romania	7	137	41	122
Slovakia	10	70	27	63
Slovenia	7	26	21	22
Spain	9	180	125	100
Sweden	8	144	99	113
United Kingdom	24	544	442	89
Total	224[a]	3862	2368	2626

[a] This number is larger than the number of selected parties in table 5 because it includes 26 parties in STV countries (Ireland, Malta and the Northern Ireland constituency in the UK).

book accounts. Christian Democrat and right-wing party candidates prefer Facebook over Twitter. All other party families do not have a clear preference, with an adoption pattern reflecting the overall number of slightly fewer Twitter accounts.

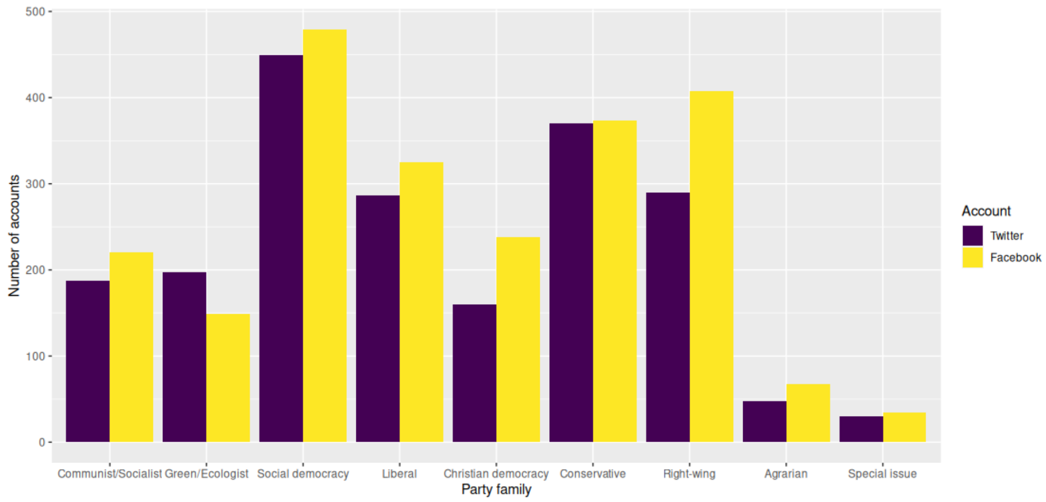


Figure 8: Twitter and Facebook accounts by party family.

17.2 Re-analysis of Nulty et al. (2016)

Another potential application of this dataset is the re-analysis of findings from earlier EP elections. Such analyses can go beyond robustness checks of previous research to inform the formation of new hypotheses which incorporate several election cycles. As an example, I re-analysed a mixed-effects logistic regression model of Twitter adoption during the 2014 EP elections by Nulty et al. (2016, Table 2) and extended it to Facebook accounts. The specification is not supposed to answer a specific research question, nor does it exactly replicate the model in Nulty et al. (2016), but it shows how substantive hypotheses can be tested with this dataset.

At the country level, the model includes internet usage, which is readily available in the dataset from Eurobarometer data. At the party level, there are two variables of a party's performance at the domestic level, specifically whether a party is in government and its vote share in the latest national election cycle. The other party-level variables capture policy positions based on the (general) left–right and liberty–authority (also called Green/Alternative/Libertarian vs. Traditionalist/Authoritarian/Nationalist) scales and attitudes towards the EU.

As described above, the dataset includes these variables by building on existing data from ParlGov (Döring et al. 2020) on parties' domestic performance, as well as data from the 2019 Chapel Hill Expert Survey (Jolly et al. 2022) on policy positions.

Turning to the candidate level, the model includes incumbency and the candidate's gender. In addition to the indicator variable for incumbents, the EuroParl identifier could be used to add more information about an incumbent's tenure in the European Parliament, such as policy positions from roll call votes or speeches (e.g. Poole and Rosenthal 1985; Lauderdale and Herzog 2016). Finally, I add the (log of the) position on the ballot paper as a new variable—which does not feature in the work by Nulty et al. (2016)—as a potential proxy variable for the likelihood of being elected independent of how many seats the party wins.

The results for Twitter accounts in Table 7 are broadly similar to the results in Nulty et al. (2016). All explanatory variables, except whether a party is part of the governing coalition at the national level, are significant at standard levels. An increase in internet usage by one percentage point is associated with 6.2% higher odds of having a Twitter account, which is substantively the same result as in Nulty et al. (2016). National party vote share is significant, but the effect size is almost negligible. The coefficients for party ideological positions show significant effects of moderate size. While the positions towards liberty/authority and EU integration are in the same direction as in Nulty et al. (2016), the general left–right coefficient is positive.[33] Regarding individual-level variables, the effects of incumbency and list position are significant and large. Incumbents have 4.5 times higher odds of using Twitter compared to non-incumbents, while a one-unit increase on the log scale of the list position, for example, a change from second to sixth position, is associated with 55% lower odds of having a Twitter account. Finally, male candidates more often use Twitter than female candidates.

The adoption of Facebook follows a completely different pattern. The only significant effects are party position towards the EU, incumbency and position on the list. The substantive size of these effects is largely comparable to the corresponding effect size in the Twitter model. The parameter estimate for internet usage is not longer significant, which is

33. However, this coefficient captures the (general) left/right effect after accounting for liberty/authority and EU positions. Suppose the left/right scale represents a combination of an economic dimension (state vs. market) and a social dimension (liberty vs. authority). In that case, the remaining variation in the left/right variable might be closer to the economic dimension which could explain this particular result.

Table 7: Mixed-Effects Logistic Regression Model for having a Twitter/Facebook account.

	Twitter	Facebook
(Intercept)	1.51^{***}	2.22^{***}
	(0.36)	(0.48)
Internet Usage (%)[a]	0.06^{***}	-0.01
	(0.01)	(0.02)
Party in Government	-0.27	0.03
	(0.18)	(0.26)
Vote Share	0.02^{*}	0.01
	(0.01)	(0.01)
Left–Right	0.09^{*}	0.05
	(0.05)	(0.07)
Liberty–Authority	-0.17^{**}	-0.12
	(0.05)	(0.08)
Anti–Pro EU	0.10^{**}	0.12^{**}
	(0.03)	(0.04)
Incumbent	1.51^{***}	1.49^{***}
	(0.19)	(0.21)
Gender: Male	0.17^{*}	0.18
	(0.09)	(0.09)
log(List Position)	-0.82^{***}	-0.59^{***}
	(0.07)	(0.08)
Random Effects (Variance):		
Country (Intercept)	0.59	0.99
Party (Intercept)	0.32	0.87
Candidates	3365	3365
Parties	177	177
Countries	28	28
AIC	3644.98	3172.70
BIC	3718.44	3246.15
Log Likelihood	-1810.49	-1574.35

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

[a] Centred at the grand mean.

a surprising result because internet usage has consistently been found to be associated with the adoption of social media tools (e.g. Nulty et al. 2016). However, the descriptive overview above already hinted at that. Table 6 shows that the number of Facebook accounts is consistently higher in most countries, apart from a few outliers.

While this re-analysis does not answer a substantive research question in itself, it shows how the dataset introduced in this paper can be used to study social media adoption in general. The juxtaposition of Twitter and Facebook is especially interesting and can shed light on differences between the two platforms. Such comparisons are only possible because accounts from both social networking sites were collected following the exact same methodology.

17.3 Further Applications

The 2019 European Parliament elections have been a fruitful ground for studying populism and populist language, often within social media communication (e.g. Alonso-Muñoz and Casero-Ripollés 2020; Fenoll 2022; Klinger et al. 2022; Carral et al. 2023). Many studies examine party accounts, only focus on a single or a few European countries, and use either Twitter or Facebook, or a combination thereof. Extending these studies beyond the limits of a given context can generate valuable insights regarding the generalisability of their findings and open up new dimensions. For example, analysing the messages of candidates in addition to party messages could shed light on whether candidates try to moderate extreme, populist language at the party level or whether they reinforce it.

18 Conclusion

This paper introduces a comprehensive dataset of candidates' social media accounts during the 2019 European Parliament elections and documents the data collection process. Starting from a complete list of all 15,540 candidates from 560 parties in 28 European countries, I selected 3,862 candidates from 198 parties for collection. Each candidate was checked and their Twitter and/or Facebook account collected, creating a dataset of 2,368 Twitter and 2,626 Facebook accounts. The dataset was expanded with additional information such as gender, party-level data, incumbency status, internet use and vote counts.

I show the usefulness of the dataset for substantive research with a descriptive overview of social media adoption patterns, followed by a mixed-effects logistic regression model which re-estimates a model by Nulty et al. (2016) based on data from the 2014 European Parliament elections. This re-analysis largely confirms earlier findings for Twitter accounts and reveals interesting patterns for Facebook accounts, which underlines the need for a comprehensive dataset to understand the phenomenon of social media adoption. Finally, I describe other research areas in which the dataset could be used to tackle new research questions and to make existing analyses more robust.

Chapter VII

(Paper 4) Cultivating Personal Votes on Social Networking Sites: An Analysis of the 2019 European Parliament elections

Abstract

Social networking sites, such as Facebook and Twitter, tend to be where citizens get news updates and information about political parties and candidates. Therefore, these sites have become an essential strategic communication instrument for parties and election campaigns to spread information and mobilise potential voters. Previous theoretical and empirical work suggests that candidates' campaign strategies are influenced by the electoral rules by which they compete because these rules incentivise either seeking personal votes or votes for the party. Consequently, there should be a connection between the electoral rules and the messaging behaviour of political candidates on social networking sites. This paper moves beyond analysing the *frequency* of campaign messages on social networking sites to consider their *style*. Using Twitter messages during the 2019 European Parliament elections, a state-of-the-art multilingual classification algorithm based on recent innovations in large language models classified tweets into *engaging* and *broadcasting* styles. Findings show that electoral system incentives do not influence the frequency of tweets. However, they relate to messaging style in that systems favouring personal vote-seeking are associated with more engaging tweets. These results hold in the subset of English, French and German language tweets, which the classifier is trained on, but do not generalise to a multilingual analysis. Future work is needed to establish whether this is due to poor classifier performance in the multilingual setting or genuine differences in the underlying data.

19 Introduction

Social networking sites—Facebook and Twitter in particular—have become important strategic tools for political campaigning. Their reasonably widespread adoption by citizens and candidates in recent years has not only made campaign strategists place increasing importance on these tools (Kreiss et al. 2017), but has also allowed researchers to study how they are used for strategic communication (Norris 2000; Chadwick 2017). The comparatively easy access to campaign messages posted on social networking sites, and their canonical format, enables researchers to conduct large-scale comparative studies of strategic communication during election campaigns using behavioural data (Jungherr and Theocharis 2017). This allows researchers to test theories about the determinants of strategic communication by directly observing and analysing communication patterns.

In this project, I used social networking sites to test how electoral rules shape candidates' incentives (Shugart 2005), especially the incentive to campaign for either candidate or party votes (Carey and Shugart 1995). This follows recent empirical work by Obholzer and Daniel (2016), who show that candidates in electoral systems with higher incentives for personal votes post more messages on Twitter. I expand on this research by not only considering the *volume* but also the *content*—specifically the messaging style—of messages.

20 Literature Review

20.1 Election Campaigns and Social Networking Sites

The use of social networking sites in general, and Twitter in particular, during election campaigns has been investigated using a multitude of theoretical and methodological approaches, even though such services have only recently become widely used (Jungherr 2015; Boulianne 2015). While much of the literature on campaigning on Twitter is focussed on national or sub-national elections (Jungherr 2015), European Parliament elections have also been widely researched. Nulty et al. (2016) provide a comprehensive overview of candidates' use of Twitter during the 2014 European elections, which includes analyses of which candidates adopted Twitter, how often they posted messages, what they talked about, and the emotional tone of their messages. Sandberg and Öhberg (2017) focus on gender differences in the adoption, use and assessment of usefulness of Twitter by Swedish candidates in the 2014 European Parliament elections. Theocharis et al. (2016) show that politicians using Twitter primarily for the purpose of broadcasting messages during the 2014 European Parliament election campaign can be attributed, in part, to having been subjected to uncivil reactions from citizens.

Despite its popularity as an object of study, as well as the important role ascribed to it by campaign professionals (Kreiss et al. 2017; Klinger and Russmann 2017), there is very little causally identified research on the effect of Twitter messages on electorally relevant outcomes (Boulianne 2015, 534). Lee (2013) found that, compared to watching a TV debate, Twitter messages provide a higher level of intimacy and, through this, a favourable view of the candidate but only for participants with a high *need for cognition*. For participants with a low *need for cognition*, the opposite is true. Those participants ascribe a higher level of intimacy, and thereby positive evaluations of the candidate, to TV debates. The second exception is Kobayashi and Ichifuji (2015), who encouraged participants to follow a Japanese gubernatorial candidate. They found that exposure to the candidate's tweets increased positive feelings towards the candidate but had no significant effect on participants' knowledge of the candidate's position on issues, evaluation of the candidate's personal traits, or voting behaviour. As excellently discussed and further investigated in a laboratory experiment, the findings are inconclusive partly due to problems involved in measuring compliance with the experimental protocol, specifically whether participants actually read the candidate's tweets (Kobayashi and Ichifuji 2015, 585–588).

20.2 Election campaigns and personal votes

Over the last 30 years, comparative election systems research, which is the study of how electoral rules shape the incentives of politicians and voters and thereby policy outcomes, has turned into a mature, well-established subfield of political science (Shugart 2005). Electoral systems have been shown to affect the incentives of parties through the distribution of votes (Taagepera and Shugart 1989; Gallagher and Mitchell 2005). European Parliament elections have been an especially fruitful research object, with empirical studies looking at politician-voter interactions mediated by the electoral system (Bowler and Farrell 1993; Sudulich et al. 2013). An especially interesting branch of comparative election systems research focuses on the link between electoral rules and candidate incentives (Carey and Shugart 1995; Shugart 2005). It posits that candidates have different incentives to solicit either personal or party votes depending on the institutional setting defined by the voting system and the district size. Candidates in systems that allow preferential voting, such as single-transferable vote or open-list proportional representation, have a stronger incentive to campaign for personal votes the larger the size of the district (Carey and Shugart 1995; Shugart 2005). For candidates competing in systems without preferential voting, such as closed-list proportional representation, larger district sizes actually decrease the incentive to campaign for personal votes (Carey and Shugart 1995; Shugart 2005). This theoretical link has been empirically studied several times in the context of European Parliament elections (Bowler and Farrell 1993, 2011). One of the most recent contributions to this question, by Obholzer and Daniel (2016), examines whether there are electoral system determinants of political campaigning on the social media platform, Twitter. They found that electoral systems were related to the message frequency of political candidates during the 2014 European Parliament elections.

20.3 Research question

The preceding discussion of the literature—especially previous research on the connection between electoral rules and campaign activities in general (Carey and Shugart 1995; Shugart 2005; Bowler and Farrell 2011) and specifically on social media (Obholzer and Daniel 2016)—led me adopt the following research question for this project:

RQ Can the incentive to cultivate a personal vote from the electoral system explain the frequency and content of candidates' social network use during election campaigns?

21 Case selection

I used the 2019 European Parliament elections to investigate the research question. European Parliament elections are generally considered second-order elections in which voters mainly reward or punish national political parties for their performance on national political issues (Reif and Schmitt 1980; Hobolt and Wittrock 2011; Hix and Marsh 2011) or use the opportunity of a "low-key" election for protest or extreme voting (Hix and Marsh 2007; Hobolt and Spoon 2012). This evidence suggests that European Parliament elections are qualitatively different from national and sub-national elections. At the same time, it stands to reason that "elections – even second-order ones – may not matter to voters, but they surely matter to candidates" (Farrell and Scully 2005, 672). This makes the study of candidate behaviour and campaign styles during European Parliament elections a worthwhile object of study to increase scientific knowledge about campaigning and its determinates.

The main advantage of using the European Parliament elections is that, although they are held under rules defined by each country individually, there are certain restrictions mandated by European Union law. The restrictions in terms of electoral rules were introduced in 2002 and mandate the election to be "on the basis of proportional representation, using the list system or the single transferable vote" (2002/772/EC, Euratom; Art. 1). This still allows variation in the type of proportional representation and the size of electoral districts, which leads to a range of actual implementations in different EU member states (Farrell and Scully 2005). According to the work of Carey and Shugart (1995) and Shugart (2005), those different rules imply different incentives for individual candidates to cultivate personal votes. As depicted in Figure 9, larger district sizes lead to stronger incentives to cultivate a personal vote in open-list proportional representation, whereas larger district sizes have the opposite effect of reducing incentives for personal votes in closed-list systems.

At the same time, the mandated proportional representation system limits the variance in the explanatory variable. However, I argue that the advantage of having parallel elections for the same political arena in a set of, at least globally speaking, broadly similar countries with a common set of laws and regulations at the EU level makes comparison between the selected countries more believable. This point is especially important because the identification of the causal effect of the electoral system on the outcome depends on the selection-on-observables assumption at the country level, as discussed in more detail below. That is, the model has to account for all country-level differences in the outcome variables. This is much more likely, a priori, and can be achieved with the inclusion of fewer control variables,

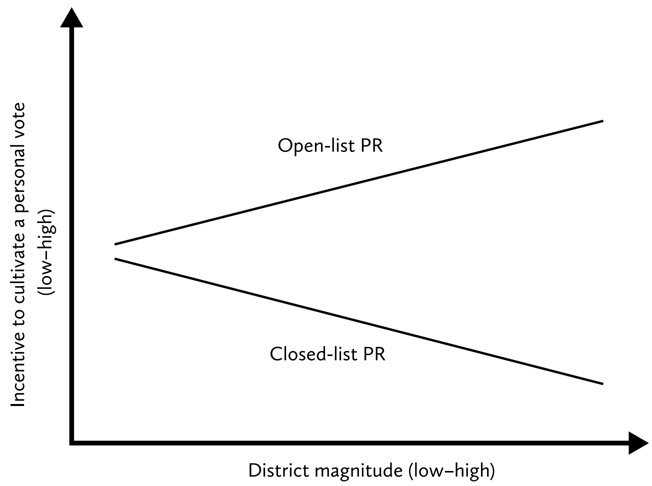


Figure 9: The effect of district magnitude on incentives to cultivate a personal vote according to Carey and Shugart (1995, 431) and Shugart (2005, 47).

a posteriori, if the countries in the sample are more similar.

21.1 Hypotheses

The first hypothesis is one of the hypotheses of Obholzer and Daniel (2016) and Daniel and Obholzer (2020). In their analysis of the 2014 European Parliament elections, they found support for this hypothesis, while their analysis of the 2019 European Parliament elections shows mixed results with no clear pattern. Both analyses focus only on incumbents.

H_1 Candidates in electoral contexts with stronger incentives for personalised campaigns use social networking sites more frequently than candidates in less personalised campaigns.

The first hypothesis rests either on the assumption that candidates in party-centered electoral systems have less incentive to campaign in general, or on the assumption that social networking sites are especially used for, and successful in, soliciting personal votes. There is limited evidence from laboratory experiments, which supports the former conjecture but also shows that this connection is mediated by the personal attributes of voters (Lee 2013; Lee and Oh 2013). There might be differences not (only) in how intensively candidates campaign, but also *how* they structure their messages. An engaging, interactive campaign style is often associated with a more personal connection between the candidate and potential voters (e.g. Stromer-Galley 2014; Hemsley et al. 2018). The second hypothesis investigates this potential link.

H_2 Candidates in electoral contexts with stronger incentives for personalised campaigns send more engaging messages (in contrast to broadcasting) than candidates in less personal campaigns.

22 Data

Data was collected from the 2,368 Twitter accounts described in Chapter VI. Tweets were collected by regularly querying the Twitter API.[34] The database contains around four million tweets, but only about 500,000 fall into the eight-week campaign period before the elections. Table 8 contains an overview of all data at the country level.

22.1 Outcome Variables

I constructed two outcome variables from the gathered campaign messages, one for each tested hypothesis. The first hypothesis posits differences between electoral systems in the number of campaign messages sent on social networking sites. This outcome variable is readily available from the collected tweets. As each message is associated with the date when it was sent, it was possible to aggregate the messages into different periods of the campaign, specifically splitting the campaign into a "long" and a "short" campaign in line with Obholzer and Daniel (2016) and Daniel and Obholzer (2020).

The outcome variable for the second hypothesis—engaging vs. broadcasting communication style—is based on the classification of candidates' messages into the following seven categories developed by Golbeck et al. (2010) and Hemphill et al. (2013):

1. Narrating: Narration of the whereabouts of candidates
2. Messaging: Sending messages to other users, or engaging in conversations with them
3. Requesting action: Asking the public to do something for the candidate
4. Thanking: Thanking the public for their support

34. At the time of data collection, the API endpoint `GET statuses/users_timeline` allowed unfiltered access to the last 3,200 messages sent by a given Twitter account.

Table 8: Number of Accounts and Tweets at the Country-Level.

Country	Accounts	Tweets (8 weeks)	Tweets (4 weeks)
Austria	43	4877	2967
Belgium	76	9783	5373
Bulgaria	27	408	222
Croatia	21	1677	896
Cyprus	13	343	162
Czechia	62	3167	1808
Denmark	63	6182	3376
Estonia	20	963	517
Finland	55	5972	3263
France	281	97601	54160
Germany	143	16447	9800
Greece	95	9863	4925
Hungary	21	460	216
Ireland	45	25021	16288
Italy	188	21259	13423
Latvia	43	4329	2302
Lithuania	23	116	35
Luxembourg	15	528	323
Malta	20	2192	1147
Netherlands	68	10722	7099
Poland	261	63556	35121
Portugal	30	4183	2615
Romania	41	969	387
Slovakia	27	153	81
Slovenia	21	5202	2912
Spain	125	62769	37130
Sweden	99	13278	7552
United Kingdom	442	199905	127762
Total	2368	571925	341862

5. Positioning: Statements about candidates' policy positions
6. Directing to information: Sharing external sources of information
7. Unrelated/other: No relation to the elections or no category applies

Categories 1–4 count as *engaging, personal* messaging styles, whereas categories 5 and 6 are indicators of a *broadcasting, general* style.

The large number of tweets (see table 8) rendered manual annotation unfeasible; therefore, a state-of-the-art supervised classification model was used to generalise an initial sample of hand-annotated messages to the full selection of candidates' messages (Hastie et al. 2009, ch. 2; Hopkins and King 2010). This approach is very common in the machine-learning literature (e.g. Burger et al. 2011; Culotta et al. 2015), and it is also frequently employed in large-scale classification tasks in political science (e.g. Boulus and Dowding 2014; Peterson and Spirling 2018). An additional problem was that the messages were authored in many different languages. The overview in Table 9 shows that tweets were almost always authored in an official language of a particular country.[35] For non-English speaking countries, English-language tweets only make up a small fraction of all tweets, which is in line with similar findings during the 2014 EP elections (Nulty et al. 2016, 432). Focusing only on English tweets would, therefore, be misleading.

Large language models, such as the one powering the well-known chatbot, ChatGPT, or the more widely available BERT (Devlin et al. 2019), can solve both problems. Similar to word embeddings, e.g. word2vec (Mikolov et al. 2013) or GloVe (Pennington et al. 2014), large language models learn generic representations of words in high-dimensional space from huge corpora. This *pre-training* step forms the basis of the model, which can then be adapted to the specific classification task at hand in the *fine-tuning* step. Thus, it is possible to take advantage of general patterns while still creating a domain-specific model. There are multilingual BERT models (e.g. Conneau et al. 2019) and models pre-trained on tweets in English (e.g. Nguyen et al. 2020), which could have been used here, but they only address one of the two problems. Zhang et al. (2022) describe the recent addition of the TwHIN-BERT model. This BERT-based model was trained on over seven billion tweets in 100 different languages. In ad-

35. Language identification is based on a pre-trained fastText model (Joulin, Grave, Bojanowski and Mikolov 2016; Joulin, Grave, Bojanowski, Douze et al. 2016).

Table 9: Distribution of languages of original tweets (without retweets and quoted tweets) based on an eight-week-long campaign.

		Language (%)			
Country	Accounts	Official	English	Other	Tweets
Austria	43	87%	6%	7%	1870
Belgium	76	65%	25%	10%	4025
Bulgaria	27	41%	38%	21%	340
Croatia	21	65%	4%	31%	951
Cyprus	13	98%	1%	1%	162
Czechia	62	87%	7%	6%	2349
Denmark	63	85%	6%	9%	3134
Estonia	20	72%	21%	7%	492
Finland	55	90%	4%	6%	3222
France	281	95%	1%	4%	27297
Germany	143	84%	9%	7%	7111
Greece	95	88%	1%	11%	6899
Hungary	21	45%	47%	8%	341
Ireland	45	94%	—	6%	9932
Italy	188	95%	1%	4%	14199
Latvia	43	84%	10%	6%	1997
Lithuania	23	48%	43%	9%	103
Luxembourg	15	75%	13%	12%	246
Malta	20	66%	—	34%	1125
Netherlands	68	78%	11%	11%	3868
Poland	261	92%	2%	6%	28162
Portugal	30	84%	4%	12%	2283
Romania	41	49%	13%	38%	774
Slovakia	27	55%	27%	18%	113
Slovenia	21	78%	4%	18%	1391
Spain	125	86%	2%	12%	13047
Sweden	99	88%	5%	7%	8086
United Kingdom	442	94%	—	6%	71298
Total	2368	90%	3%	7%	214817

dition to the tweet text, the authors also include information from the social interaction on Twitter, for example, favourites, replies or retweets, which helps the model to learn patterns that would be hard to detect from the text alone (Zhang et al. 2022).

To generate training data for the *fine-tuning* of the TwHIN-BERT-based classifier, I decided to select languages that are spoken in more than one country. English was the first selection because it is an official language in the UK, the Republic of Ireland and Malta. I then selected German, which is an official language in Germany, Austria and Belgium, and one the commonly spoken languages in Luxembourg. Finally, I selected French, which is spoken in France and Belgium and is also commonly spoken in Luxembourg. After randomly selecting 1,200 English and German, and 400 French tweets, I used a combination of crowd-coders (K. Benoit et al. 2016; Lehmann and Zobel 2018) and two research assistants to classify the randomly selected tweets into one of the seven categories described at the beginning of this section (see Section E.1 in the Appendix for the codebook). The classification was implemented as a survey on Qualtrics, while crowd-coders were recruited through Prolific.[36] After discarding tweets with low coder agreement, the training data contained 1,073, 916 and 386 tweets in English, German and French, respectively.

Table 10: Tweet Classification Model Overview.

	Model EN/DE	Model EN/DE/FR
Languages	English, German	English, German, French
Training samples	1989	2375
Train-test split	80/20	80/20
Training epoch	3	3
Learning rate	5×10^{-5}	3×10^{-5}
(Weighted) Precision	0.827	0.816
(Weighted) Recall	0.771	0.696
(Weighted) F1 Score	0.764	0.748

To train the classification model, I first aggregated *narrating*, *messaging*, *requesting action* and *thanking* into *engaging* and *positioning* and *directing to information* into *broadcasting*, while keeping *unrelated/other* as is.[37] Then, I trained two classifiers: one using just the

36. `https://www.prolific.com/`; Tasks were advertised at £9/h, with effective compensation between £10.50 and £16.10.

37. The alternative strategy of training the classification model on all categories and aggregating afterwards

English and German training data because most of the training data is in these languages, and the other using all three languages to allow the classifier to learn to transfer between several languages. The fine-tuning was based on the TwHIN-BERT base model described above. This process can be adjusted by many parameters, but the most important are the number of training epochs and the (initial) learning rate for the stochastic gradient descent algorithm, specifically AdamW (Loshchilov and Hutter 2017). The training epochs parameter controls how many passes the algorithm makes through the whole dataset, which roughly corresponds to the number of "iterations". The learning rate, or step size, controls how fast the algorithm changes parameters. For fine-tuning BERT models, Devlin et al. (2019, 4183–4184) recommend sensible values of 5×10^{-5}, 3×10^{-5} and 2×10^{-5} for the learning rate, and 2, 3 and 4 for the number of epochs. A grid search on these parameters yielded the configuration shown in Table 10 with the best models having a (class-weighted) F1 score of 0.764 for the English/German and 0.748 for the English/German/French model.

22.2 Explanatory and Control Variables

Following the theory of how different combinations of election rules and district magnitudes can influence campaign style, I use variables directly corresponding to each concept. The *open list* variable indicates whether the electoral system is closed or open list. In this context, Single Transferable Vote systems are counted as open list systems. The second variable measures the district size at the country level by averaging over the number of eligible voters per district. There are two countries which required special processing. Firstly, Belgium does not release statistics about overall voters at the level of the Dutch, French and German electoral colleges (districts) for political reasons. Therefore, Belgium is counted as one big district instead of the (actual) average over three smaller districts. Secondly, Germany nominally uses states as electoral districts, but only the "union parties" CDU/CSU use lists at the state level. All other parties use a single, national list; thus, Germany is counted as a single district.

Table 11 gives an overview of control variables which have been used in previous work (e.g. Nulty et al. 2016; Obholzer and Daniel 2016; Daniel and Obholzer 2020; Bowler and Farrell 2011). The column *Used* shows which ones are included in the models estimated below. While

performed worse in this case.

Table 11: Overview of explanatory and control variables used in previous research.

Level	Variable	Type	Source[a]	Used
Country	Preferential voting	Explanatory	OD	✓
	Avg. District Magnitude	Explanatory	OD	✓
	Avg. Citizens Represented	Explanatory	OD	✓
	Internet usage	Control	NA	✓
	Social network usage	Control	OD	
National Party	Extreme (left-right)	Control	OD	✓
	Extreme (GAL-TAN)	Control	OD	✓
	Extreme (EU Integration)	Control	OD	✓
	Seat Share (national)	Control	OD	✓
	In Government (national)	Control	OD	✓
Candidate	Gender	Control	OD, NA	✓
	Age	Control	OD, NA	
	Incumbent	Control	OD, NA	✓
	European Party Leader	Control	OD	
	EP Committee Leader	Control	OD	
	EP Spitzenkandidat	Control	OD	✓
	EU position	Control	NA	
	Left-right position	Control	NA	
	List position safety	Control	OD	
	Twitter Followers	Control	OD	✓
	Tweets	Control	—	✓

[a] OD refers to Obholzer and Daniel (2016) and NA to Nulty et al. (2016).

some candidate-level variables are not available for non-incumbents, the included variables still follow best practice in the literature on the effects of electoral systems and on the use of social networking sites in European Parliament elections.

23 Model Specification

The data has a multilevel structure. Messages on Twitter are authored by a candidate, who is a member of a national party, which is again nested in a country. Both hypotheses posit effects of the electoral system, defined separately in each country, on different outcome variables. The main explanatory variables are at the country level. Individual messages from Twitter are at the lowest level. For hypothesis 1, those messages are aggregated by candidate such that candidates become the lowest level while the outcome variable for the second hypothesis is observed at the level of the message. Table 12 gives an overview of the model specification. All models are specified with random intercepts at the country and party level, and are estimated with the *lme4* package (Bates et al. 2015) in R.[38]

Table 12: Multilevel model specification by hypothesis.

	H_1	H_2
Outcome	Number of Messages	Engaging/Broadcasting
Observation level	Candidate	Message
Measure Level	Count	Binary
Model	Neg. Binom.	Logistic

24 Results

Table 13 shows the results of testing the first hypothesis with four different specifications. Separate models were estimated for the two- and one-month-long campaigns and a sample comprising all candiates and a sample comprising incumbents only. The only consistently significant effects are that the more tweets a candidate has written prior the election campaign, the more tweets they will post during the campaign, and that incumbents send fewer

38. Additional packages: *texreg* (Leifeld 2013) and *ggplot2* (Wickham 2016).

Table 13: Mixed-Effects Negative Binomial Model for the Number of Tweets (Outcome) based on all Tweets.

	Two-month campaign		One-month campaign	
	All Candidates	Incumbents	All Candidates	Incumbents
(Intercept)	1.23***	0.36	0.89**	−0.11
	(0.24)	(0.44)	(0.27)	(0.51)
Internet Usage[a]	0.01	0.01	0.02*	0.02
	(0.01)	(0.01)	(0.01)	(0.01)
Citizens Represented[b]	0.25	0.26*	0.32	0.37*
	(0.14)	(0.13)	(0.16)	(0.15)
Open List	−0.55	−0.47	−0.56	−0.53
	(0.42)	(0.35)	(0.50)	(0.42)
District Size[b]	−0.14	−0.07	−0.17	−0.11
	(0.15)	(0.12)	(0.18)	(0.14)
Open List × District Size[b]	−0.14	−0.82	−0.12	−1.08
	(0.73)	(0.66)	(0.88)	(0.78)
In Government	−0.14	−0.19	−0.17*	−0.23
	(0.08)	(0.13)	(0.08)	(0.14)
Vote Share	−0.00	0.00	−0.00	0.00
	(0.00)	(0.01)	(0.00)	(0.01)
Left–Right	0.02	0.03	0.05	0.04
	(0.02)	(0.04)	(0.03)	(0.05)
GAL–TAN	−0.02	−0.05	−0.04	−0.06
	(0.03)	(0.04)	(0.03)	(0.05)
Anti–Pro EU	0.01	−0.01	0.03	−0.00
	(0.01)	(0.03)	(0.01)	(0.03)
Gender: Male	−0.09*	−0.13	−0.10*	−0.12
	(0.04)	(0.09)	(0.05)	(0.11)
EP Spitzenkandidat	0.46	0.27	0.51	0.40
	(0.34)	(0.35)	(0.37)	(0.40)
log(Followers)	−0.01	−0.05	−0.00	−0.05
	(0.02)	(0.04)	(0.02)	(0.05)
log(Tweets)	0.54***	0.66***	0.49***	0.64***
	(0.02)	(0.05)	(0.02)	(0.06)
Incumbent	−0.20**		−0.17*	
	(0.06)		(0.07)	
Random Effects (Variance):				
Country (Intercept)	0.19	0.08	0.29	0.13
Party (Intercept)	0.04	0.05	0.04	0.05
Candidates	1791	321	1791	321
Parties	163	102	163	102
Countries	28	28	28	28
Dispersion Parameter	1.37	1.77	1.12	1.29
AIC	21145.44	3902.90	19414.37	3598.01
BIC	21249.76	3970.79	19518.69	3665.89
Log Likelihood	−10553.72	−1933.45	−9688.19	−1781.00

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; [a] Centred at the overall (grand) mean; [b] Standardised.

tweets than non-incumbents. Additionally, male candidates use Twitter less often than female candidates in the analysis using the sample of all candidates. Comparing this with the adoption of Twitter accounts in Chapter VI (Table 7) suggests that while male candidates are more likely to have a Twitter account, they use it less often to send tweets. Finally, the more citizens a candidate represents, the more they tweet, but only for the incumbent sample. However, also note that this variable is standardised.

These results do not support the first hypothesis. Electoral system incentives do not have an influence on how often candidates use Twitter. While this contradicts findings from the elections in 2014 (Obholzer and Daniel 2016), it is consistent with the re-analysis of the 2019 data by the same authors (Daniel and Obholzer 2020). The results from the incumbent sample in Table 13 are not directly comparable to the results of Daniel and Obholzer (2020, Table 1) because they use a linear instead of a count model, and they include different control variables. Limiting the sample to incumbents gives them access to additional candidate-level control variables such as age and years of service in the EU Parliament. Despite these differences in the specification of the model, Daniel and Obholzer (2020, 3) reach the same conclusions that the electoral system incentives do work in the same way as they did in 2014. As a robustness check, I also ran an alternative model based on original tweets only, that is without retweets and quoted tweets. The results in Table 31 in the Appendix lead to the same conclusions about the first hypothesis.

To test the second hypothesis, I estimated three models. Model 1 includes English and German tweets with classifications from the *EN/DE* model, which was trained on English and German tweets. Similarly, model 2 includes tweets in English, German and French with output from the *EN/DE/FR* classification model. Predicting labels for tweets in the same language(s) on which the classifier was trained is advantageous because the performance of the classifier is known (see Table 10). Unfortunately, it limits the sample to English, German and French tweets, only three out of the 24 official languages of the EU. To address this limitation, I estimated a third statistical model using tweets in all languages based on the *EN/DE/FR* classifier. I expect this classification model to perform best because it is based on three languages, which should make it easier to transfer the learning to the other 21 languages in the sample. However, the actual performance is unknown because the hand-annotated data only exists for English, German and French. The results from the third model should consequently be treated cautiously.

Table 14 shows the estimates of the three models testing the second hypothesis. The results from the first two models are largely similar in significance and effect size. In both models, higher internet usage is associated with more engaging tweets, and the more citizens a candidate represents, the more they broadcast rather than engage. The key theoretical explanatory variables show a pattern that closely follows the theoretical expectations in Figure 9. In the first model, increasing district size by one standard deviation is associated with a decrease in the odds of an engaging vs a broadcasting messaging style by 21.3% in closed-list systems, whereas in open-list systems, the same change increases the odds of more engaging messages by almost 400%. The latter effect remains significant in model 2, while the former effect loses its significance, suggesting that district size has no effect for candidates in closed-list systems when including French-language tweets.

While parties' ideological positions towards the EU influence campaign style, candidates from cabinet parties are associated with a more broadcasting style. Vote share has a significant positive effect but with a very small effect size because vote share is measured as a proportion between 0 and 1. At the candidate level, gender, incumbency and whether the candidate is an EP Spitzenkandidat are significant. A consistent finding in all three models is that female candidates adopt a more engaging message style compared to male candidates. This result extends the research of Sandberg and Öhberg (2017), who found that female candidates place more importance on personal campaign activities than their male counterparts. Incumbents also consistently send more broadcasting messages. Finally, Spitzenkandidaten are associated with a less personal messaging style in models 1 and 2, but this effect is completely reversed in model 3.

Overall, the evidence for H_2 is mixed. The findings in models 1 and 2 lend support to the hypothesis. However, these results are mostly confined to the countries where English, German or French is an official language. Although the summary statistics show that tweets from candidates in 27 or 28 countries are included in the models, the language distribution in Table 9 makes it clear that the bulk of the data comes from a small set of countries, specifically the UK, Ireland, Germany, Austria, France, Belgium and Luxembourg. None of those

Table 14: Mixed-Effects Logistic Regression Model for Engaging vs. Broadcasting Tweets.

	Model 1	Model 2	Model 3
(Intercept)	1.38^{***}	2.35^{***}	0.95^{**}
	(0.38)	(0.40)	(0.31)
Internet Usage[a]	0.04^{***}	0.05^{***}	0.04^{***}
	(0.01)	(0.01)	(0.01)
Citizens Represented[b]	-0.11^{*}	-0.16^{**}	0.09
	(0.05)	(0.06)	(0.15)
Open List	0.09	0.00	-0.70^{*}
	(0.27)	(0.29)	(0.31)
District Size[b]	-0.24^{*}	-0.18	-0.25
	(0.10)	(0.10)	(0.18)
Open list $\times$ District Size[b]	1.61^{**}	1.55^{**}	0.57
	(0.51)	(0.56)	(0.47)
In Government	-0.65^{**}	-0.62^{**}	-0.13
	(0.20)	(0.21)	(0.12)
Vote Share	0.02^{***}	0.02^{***}	0.02^{***}
	(0.00)	(0.00)	(0.00)
Left–Right	0.09	0.04	0.11^{*}
	(0.07)	(0.08)	(0.04)
GAL–TAN	-0.04	-0.05	-0.15^{**}
	(0.08)	(0.08)	(0.05)
Anti–Pro EU	-0.09^{*}	-0.10^{*}	0.04
	(0.04)	(0.04)	(0.03)
Gender: Male	-0.28^{***}	-0.27^{***}	-0.31^{***}
	(0.03)	(0.02)	(0.02)
EP Spitzenkandidat	-0.51^{**}	-0.36^{*}	0.44^{***}
	(0.15)	(0.16)	(0.10)
log(Followers)	0.01	0.00	-0.07^{***}
	(0.01)	(0.01)	(0.01)
log(Tweets)	0.01	-0.07^{***}	0.07^{***}
	(0.01)	(0.01)	(0.01)
Incumbent	-0.32^{***}	-0.16^{***}	-0.17^{***}
	(0.04)	(0.03)	(0.02)
Random Effects (Variance):			
Party (Intercept)	1.09	1.31	0.69
Country (Intercept)	0.00	0.00	0.25
Language (Intercept)			0.31
Tweets	51046	69145	117070
Parties	139	139	157
Countries	27	28	28
Languages	2	3	24
AIC	43902.40	67192.05	125287.74
BIC	44061.52	67356.64	125471.48
Log Likelihood	-21933.20	-33578.03	-62624.87

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; [a] Centred at the overall (grand) mean;
[b] Standardised.

countries are in Eastern, Southern or Northern Europe, so these results might apply only in Central Europe and the British Isles. In model 3, key explanatory variables are either insignificant or go in the wrong direction, as in the case of the coefficient for open lists. However, the quality of these estimates depends on whether the classifier can extend training data in three languages to 21 other languages. Given the limited training data in a multilingual context, it is unknown how well the classifier generalises to other languages, especially those outside of the Romance and Germanic language families, and languages not written in Latin script.

25 Conclusion

This paper investigates whether electoral systems, specifically open/closed lists and district size, influence how often and in what style candidates in the 2019 European Parliament elections use Twitter. The results show that electoral systems do not influence the frequency of Twitter use. At the same time, these incentives have an influence on the campaign style—engaging vs. broadcasting—for some countries and languages, but these effects do not materialise when including all official EU languages. However, other interesting patterns which have an influence on campaign style emerge, for example with respect to gender differences.

One important limitation of these results is the unknown quality of the multilingual classification of tweets into engaging and broadcasting. Therefore, it is impossible to know whether the differences between the results for English, German and French and the multilingual arise due to underlying patterns or simply because of low quality classifications. The overall performance of the classifier, which is partly based on how good the training data is and how much is available to learn from, is a limitation in itself, even in the absence of the challenges of multilingual classification.

Future work could investigate engaging/broadcasting messaging styles on other social media platforms such as Facebook and extend the analysis to additional contexts. The upcoming 2024 European Parliament elections are an especially interesting case because not all results transfer across election cycles, as Daniel and Obholzer (2020) show for the 2014 to 2019 period.

Chapter VIII

Conclusion

This thesis comprises three research projects which examine election campaign strategies from different angles. Chapter II shows that GOTV leaflets can affect the turnout of households beyond the one contacted. The presence and size of this effect depend on the composition of the targeted household and the neighbourhood. Contacted households mobilise other households from the same party when the share or rival party supporter in the neighbourhood is high. Rival party households, on the other hand, are most often mobilised in mixed-partisan neighbourhoods.

Chapter IV examines the performance of Swiss parties' social media campaigns on Twitter during the 2015 Federal election campaign. It uses the follower network as an indicator of organisational cohesion and programmatic coherence based on Twitter message content. The results suggest that the Social Democrats and smaller, newer parties have higher organisational cohesion levels than the established, centrist and right-wing parties. Most parties show comparable, low levels of programmatic coherence compared to the Social Democrats.

Chapter VI introduces a comprehensive dataset of candidates' social media accounts in the 2019 European Parliament elections. Checking 3,862 candidates for their social media presence reveals 2,368 Twitter and 2,626 Facebook users. The discussion of potential applications of the dataset re-examines a model by Nulty et al. (2016) on the adoption of social media accounts, and outlines further research areas that could benefit from this data.

Chapter VII examines whether election system incentives to seek personal votes or votes for the party influence candidates' social media use. The results for the number of tweets show that none of the electoral system variables are significant. However, the electoral system influences the messaging style. Using a multilingual classification model to detect engaging- and broadcasting-style tweets shows that electoral systems favouring personal vote-seeking are associated with more engaging tweets but only when authored in English, German or French.

Chapter IX

Bibliography

Alonso-Muñoz, Laura and Andreu Casero-Ripollés. 2020. “Populism Against Europe in Social Media: The Eurosceptic Discourse on Twitter in Spain, Italy, France, and United Kingdom During the Campaign of the 2019 European Parliament Election”. *Frontiers in Communication* 5:54.

Ausserhofer, Julian and Axel Maireder. 2013. “National Politics on Twitter”. *Information, Communication & Society* 16, no. 3 (April): 291–314.

Bafumi, Joseph and Michael C. Herron. 2010. “Leapfrog Representation and Extremism: A Study of American Voters and Their Members in Congress”. *American Political Science Review* 104 (3): 519–542.

Bailer, Stefanie. 2017. “To use the whip or not: Whether and when party group leaders use disciplinary measures to achieve voting unity”. *International Political Science Review* 39, no. 2 (January): 163–177.

Barberá, Pablo. 2015. “Birds of the Same Feather Tweet Together: Bayesian Ideal Point Estimation Using Twitter Data”. *Political Analysis* 23 (1): 76–91.

Bartels, Larry M. 1991. “Constituency Opinion and Congressional Policy Making: The Reagan Defense Buildup”. *American Political Science Review* 85, no. 2 (June): 457–474.

Bates, Douglas, Martin Mächler, Ben Bolker and Steve Walker. 2015. “Fitting Linear Mixed-Effects Models Using lme4”. *Journal of Statistical Software* 67 (1): 1–48.

Bello, Jason and Meredith Rolfe. 2014. “Is influence mightier than selection? Forging agreement in political discussion networks during a campaign”. *Social Networks* 36 (January): 134–146.

Benoit, Kenneth, Drew Conway, Benjamin E. Lauderdale, Michael Laver and Slava Mikhaylov. 2016. “Crowd-sourced Text Analysis: Reproducible and Agile Production of Political Data”. *American Political Science Review* 110, no. 2 (May): 278–295.

Benoit, William L., Mark J. Glantz, Anji L. Phillips, Leslie A. Rill, Corey B. Davis, Jayne R. Henson and Leigh Anne Sudbrock. 2011. "Staying 'On Message': Consistency in Content of Presidential Primary Campaign Messages Across Media". *American Behavioral Scientist* 55, no. 4 (February): 457–468.

Bonica, Adam. 2014. "Mapping the Ideological Marketplace". *American Journal of Political Science* 58 (2): 367–386.

Boulianne, Shelley. 2015. "Social media use and participation: a meta-analysis of current research". *Information, Communication & Society* 18, no. 5 (March): 524–538.

Boulus, Paul and Keith Dowding. 2014. "The press and issue framing in the Australian mining tax debate". *Australian Journal of Political Science* 49, no. 4 (September): 694–710.

Bowler, Shaun, David M. Farrel and Richard S. Katz. 1999. "Party cohesion, party discipline, and parliaments". In *Party discipline and parliamentary government,* edited by Shaun Bowler, David M. Farrel and Richard S. Katz, 3–22. Columbus, OH: Ohio State University Press.

Bowler, Shaun and David M. Farrell. 1993. "Legislator Shirking and Voter Monitoring: Impacts of European Parliament Electoral Systems upon Legislator-Voter Relationships". *Journal of Common Market Studies* 31, no. 1 (March): 45–70.

———. 2011. "Electoral institutions and campaigning in comparative perspective: Electioneering in European Parliament elections". *European Journal of Political Research* 50, no. 5 (March): 668–688.

Brown, Jacob R. and Ryan D. Enos. 2021. "The measurement of partisan sorting for 180 million voters". *Nature Human Behaviour* 5, no. 8 (March): 998–1008.

Burger, John D., John C. Henderson, George Kim and Guido Zarrella. 2011. "Discriminating Gender on Twitter". In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing, 27–31 July 2011, Edinburgh,* 1301–1309. Association for Computational Linguistics. http://www.aclweb.org/anthology/D11-1120.

Cardenal, Ana S. 2011. "Why mobilize support online? The paradox of party behaviour online". *Party Politics* 19, no. 1 (July): 83–103.

Carey, John M. 2007. "Competing Principals, Political Institutions, and Party Unity in Legislative Voting". *American Journal of Political Science* 51, no. 1 (January): 92–107.

Carey, John M. and Matthew Søberg Shugart. 1995. "Incentives to cultivate a personal vote: A rank ordering of electoral formulas". *Electoral Studies* 14, no. 4 (December): 417–439.

Carral, Uxía, Jorge Tuñón and Carlos Elías. 2023. "Populism, cyberdemocracy and disinformation: analysis of the social media strategies of the French extreme right in the 2014 and 2019 European elections". *Humanities and Social Sciences Communications* 10, no. 1 (January).

Chadwick, Andrew. 2017. *The Hybrid Media System: Politics and Empowerment.* 2nd ed. Oxford: Oxford University Press.

Clauset, Aaron, M. E. J. Newman and Cristopher Moore. 2004. "Finding community structure in very large networks". *Physical Review E* 70, no. 6 (December).

Clinton, Joshua D. 2012. "Using Roll Call Estimates to Test Models of Politics". *Annual Review of Political Science* 15, no. 1 (June): 79–99.

Colleoni, Elanor, Alessandro Rozza and Adam Arvidsson. 2014. "Echo Chamber or Public Sphere? Predicting Political Orientation and Measuring Political Homophily in Twitter Using Big Data". *Journal of Communication* 64 (2): 317–332.

Coman, Emanuel Emil. 2015. "Institutions and Vote Unity in Parliaments: Evidence from 33 National Chambers". *The Journal of Legislative Studies* 21, no. 3 (June): 360–389.

Conneau, Alexis, Kartikay Khandelwal, Naman Goyal, Vishrav Chaudhary, Guillaume Wenzek, Francisco Guzmán, Edouard Grave, Myle Ott, Luke Zettlemoyer and Veselin Stoyanov. 2019. "Unsupervised Cross-lingual Representation Learning at Scale" (November). arXiv: 1911.02116.

Conover, Michael D., Bruno Gonçalves, Alessandro Flammini and Filippo Menczer. 2012. "Partisan asymmetries in online political activity". *EPJ Data Science* 1, no. 6 (June): 1–19.

Costenbader, Elizabeth and Thomas W. Valente. 2003. "The stability of centrality measures when networks are sampled". *Social Networks* 25, no. 4 (October): 283–307.

Culotta, Aron, Nirmal Ravi Kumar and Jennifer Cutler. 2015. "Predicting the Demographics of Twitter Users from Website Traffic Data". In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, January 25–30, 2015, Austin, Texas, USA,* edited by Blai Bonet and Sven Koenig, 72–78. AAAI Press. ISBN: 978-1-57735-698-1.

Dahlgaard, Jens Olav, Yosef Bhatti, Jonas Hedegaard Hansen and Kasper M. Hansen. 2021. "Living Together, Voting Together: Voters Moving in Together Before an Election Have Higher Turnout". *British Journal of Political Science* 52, no. 2 (February): 631–648.

Dalton, Russel J., Ian McAllister and Martin P. Wattenberg. 2000. "The Consequences of Partisan Dealignment". In *Parties without Partisans: Political Change in Advanced Industrial Democracies,* edited by Russel J. Dalton and Martin P. Wattenberg, 37–63. Oxford: Oxford University Press.

Daniel, William T. and Lukas Obholzer. 2020. "Reaching out to the voter? Campaigning on Twitter during the 2019 European elections". *Research & Politics* 7, no. 2 (April): 1–8.

Daniel, William T., Lukas Obholzer and Steffen Hurka. 2017. "Static and dynamic incentives for Twitter usage in the European Parliament". *Party Politics* 25, no. 6 (December): 771–781.

De Vries, Erik, Martijn Schoonvelde and Gijs Schumacher. 2018. "No Longer Lost in Translation: Evidence that Google Translate Works for Comparative Bag-of-Words Text Applications". *Political Analysis* 26, no. 4 (September): 417–430.

Devlin, Jacob, Ming-Wei Chang, Kenton Lee and Kristina Toutanova. 2019. "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding". In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers),* 1:4171–4186.

Döring, Holger, Constantin Huber and Philip Manow. 2020. *Parliaments and governments database (ParlGov): Information on parties, elections and cabinets in established democracies.* https://www.parlgov.org.

Downs, Anthony. 1957. *An Economic Theory of Democracy.* Addison Wesley.

Enos, Ryan D. 2016. "What the Demolition of Public Housing Teaches Us about the Impact of Racial Threat on Political Behavior". *American Journal of Political Science* 60, no. 1 (January): 123–142.

Enos, Ryan D. 2017. *The Space between Us: Social Geography and Politics.* Cambridge: Cambridge University Press.

Esser, Frank and Jörg Matthes. 2013. "Mediatization Effects on Political News, Political Actors, Political Decisions, and Political Audiences". In *Democracy in the Age of Globalization and Mediatization,* edited by Hanspeter Kriesi, Sandra Lavenex, Frank Esser, Jörg Matthes, Marc Bühlmann and Daniel Bochsler. Palgrave Macmillan.

European Commission, Brussels. 2019. *Eurobarometer 90.3 (2018)* [in en].

Farrell, David M. and Roger Scully. 2005. "Electing the European Parliament: How Uniform are 'Uniform' Electoral Systems?" *Journal of Common Market Studies* 43, no. 5 (December): 969–984.

Fenoll, Vicente. 2022. "Political communication on Facebook and populism. The 2019 European Parliament election in Spain". *Communication & Society* 35, no. 3 (June): 91–103.

Fichter, Adrienne and Alexandra Kohler. 2015. "Wie die Parteien Social Media im Wahlkampf nutzen". *Neue Zürcher Zeitung,* 19 September 2015.

Fieldhouse, Ed, Laurence Lessard-Phillips and Bruce Edmonds. 2015. "Cascade or echo chamber? A complex agent-based simulation of voter turnout". *Party Politics* 22, no. 2 (October): 241–256.

Fieldhouse, Edward and David Cutts. 2016. "Shared Partisanship, Household Norms and Turnout: Testing a Relational Theory of Electoral Participation". *British Journal of Political Science* 48, no. 3 (May): 807–823.

Foos, Florian. 2018. "The parliamentary candidate as persuader: Evidence from randomized candidate-voter interactions". http://www.florianfoos.net/resources/Candidate_as_persuader_Foos.pdf.

Foos, Florian and Eline A. de Rooij. 2017a. "All in the Family: Partisan Disagreement and Electoral Mobilization in Intimate Networks - A Spillover Experiment". *American Journal of Political Science* 61 (2): 289–304.

———. 2017b. "The role of partisan cues in voter mobilization campaigns: Evidence from a randomized field experiment". *Electoral Studies* 45 (February): 63–74.

Fowler, James H. 2005. "Turnout in a Small World". In *The Social Logic of Politics: Personal Networks as Contexts for Political Behaviour,* edited by Alan S. Zuckerman, 269–288. Philadelphia, PA: Temple University Press.

Freelon, Deen. 2017. "Campaigns in control: Analyzing controlled interactivity and message discipline on Facebook". *Journal of Information Technology & Politics* 14, no. 2 (April): 168–181.

Freeman, Linton C. 1977. "A Set of Measures of Centrality Based on Betweenness". *Sociometry* 40, no. 1 (March): 35–41.

———. 1979. "Centrality in social networks: Conceptual clarification". *Social Networks* 1, no. 3 (January): 215–239.

Galaskiewicz, Joseph. 1991. "Estimating point centrality using different network sampling techniques". *Social Networks* 13, no. 4 (December): 347–386.

Gallagher, Michael and Paul Mitchell, eds. 2005. *The Politics of Electoral Systems.* Oxford: Oxford University Press.

Gay, Claudine. 2011. "Moving to Opportunity: The Political Effects of a Housing Mobility Experiment". *Urban Affairs Review* 48, no. 2 (November): 147–179.

Gelman, Andrew and Gary King. 1993. "Why Are American Presidential Election Campaign Polls So Variable When Votes Are So Predictable?" *British Journal of Political Science* 23, no. 4 (October): 409–451.

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

Gervais, Bryan T. 2015. "Incivility Online: Affective and Behavioral Reactions to Uncivil Political Posts in a Web-based Experiment". *Journal of Information Technology & Politics* 12, no. 2 (January): 167–185.

Gibson, James L., Cornelius P. Cotter, John F. Bibby and Robert J. Huckshorn. 1983. "Assessing Party Organizational Strength". *American Journal of Political Science* 27, no. 2 (May): 193.

Gibson, Rachel and Andrea Römmele. 2001. “Changing Campaign Communications: A Party-Centered Theory of Professionalized Campaigning”. *Harvard International Journal of Press/Politics* 6, no. 4 (September): 31–43.

Gibson, Rachel K. 2015. “Party change, social media and the rise of ‘citizen-initiated’ campaigning”. *Party Politics* 21, no. 2 (March): 183–197.

Golbeck, Jennifer, Justin M. Grimes and Anthony Rogers. 2010. “Twitter use by the U.S. Congress”. *Journal of the American Society for Information Science and Technology* 61 (8): 1612–1621.

Graham, Todd, Marcel Broersma, Karin Hazelhoff and Guido van ’t Haar. 2013. “Between Broadcasting Political Messages and Interacting with Voters: The use of Twitter during the 2010 UK general election campaign”. *Information, Communication & Society* 16 (5): 692–716.

Grant, Will J., Brenda Moon and Janie Busby Grant. 2010. “Digital Dialogue? Australian Politicians’ use of the Social Network Tool Twitter”. *Australian Journal of Political Science* 45, no. 4 (December): 579–604.

Hainmueller, Jens, Jonathan Mummolo and Yiqing Xu. 2019. “How Much Should We Trust Estimates from Multiplicative Interaction Models? Simple Tools to Improve Empirical Practice”. *Political Analysis* 27, no. 2 (April): 163–192.

Hastie, Trevor, Robert Tibshirani and Jerome Friedman. 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* New York: Springer.

Hellwig, Timothy and David Samuels. 2007. “Voting in Open Economies: The Electoral Consequences of Globalization”. *Comparative Political Studies* 40, no. 3 (March): 283–306.

Hemphill, Libby, Jahna Otterbacher and Matthew Shapiro. 2013. “What’s congress doing on Twitter?” In *Proceedings of the 2013 conference on Computer supported cooperative work,* 877–886. New York, NY: Association for Computing Machinery (ACM).

Hemsley, Jeff, Jennifer Stromer-Galley, Bryan Semaan and Sikana Tanupabrungsun. 2018. “Tweeting to the Target: Candidates’ Use of Strategic Messages and @Mentions on Twitter”. *Journal of Information Technology & Politics* 15, no. 1 (January): 3–18.

Himelboim, Itai, Derek Hansen and Anne Bowser. 2013. "Playing in the same Twitter network: Political information seeking in the 2010 US gubernatorial elections". *Information, Communication & Society* 16, no. 9 (November): 1373–1396.

Hix, Simon. 2004. "Electoral Institutions and Legislative Behavior: Explaining Voting Defection in the European Parliament". *World Politics* 56, no. 2 (January): 194–223.

Hix, Simon and Michael Marsh. 2007. "Punishment or Protest? Understanding European Parliament Elections". *The Journal of Politics* 69, no. 2 (May): 495–510.

———. 2011. "Second-order effects plus pan-European political swings: An analysis of European Parliament elections across time". *Electoral Studies* 30, no. 1 (March): 4–15.

Hobolt, Sara B. and Jae-jae Spoon. 2012. "Motivating the European voter: Parties, issues and campaigns in European Parliament elections". *European Journal of Political Research* 51, no. 6 (April): 701–727.

Hobolt, Sara B. and Jill Wittrock. 2011. "The second-order election model revisited: An experimental test of vote choices in European Parliament elections". *Electoral Studies* 30, no. 1 (March): 29–40.

Höglinger, Dominic. 2008. "Verschafft die direkte Demokratie den Benachteiligten mehr Gehör? Der Einfluss institutioneller Rahmenbedingungen auf die mediale Präsenz politischer Akteure". *Swiss Political Science Review* 14, no. 2 (June): 207–243.

Holbrook, Thomas M. and Scott D. McClurg. 2005. "The Mobilization of Core Supporters: Campaigns, Turnout, and Electoral Composition in United States Presidential Elections". *American Journal of Political Science* 49, no. 4 (October): 689–703.

Hopkins, Daniel J. and Gary King. 2010. "A Method of Automated Nonparametric Content Analysis for Social Science". *American Journal of Political Science* 54 (1): 229–247.

Huckfeldt, Robert, Eric Plutzer and John Sprague. 1993. "Alternative Contexts of Political Behavior: Churches, Neighborhoods, and Individuals". *The Journal of Politics* 55, no. 2 (May): 365–381.

Huckfeldt, Robert and John Sprague. 1987. "Networks in Context: The Social Flow of Political Information". *American Political Science Review* 81, no. 4 (December): 1197–1216.

Huckfeldt, Robert and John Sprague. 1992. “Political Parties and Electoral Mobilization: Political Structure, Social Structure, and the Party Canvass”. *American Political Science Review* 86, no. 1 (March): 70–86.

———. 1995. *Citizens, Politics, and Social Communication: Information and Influence in an Election Campaign.* Cambridge: Cambridge University Press. ISBN: 9780521452984.

Jolly, Seth, Ryan Bakker, Liesbet Hooghe, Gary Marks, Jonathan Polk, Jan Rovny, Marco Steenbergen and Milada Anna Vachudova. 2022. “Chapel Hill Expert Survey trend file, 1999–2019”. *Electoral Studies* 75 (February): 102420.

Joulin, Armand, Edouard Grave, Piotr Bojanowski, Matthijs Douze, Hérve Jégou and Tomas Mikolov. 2016. “FastText.zip: Compressing text classification models”, arXiv: 1612.03651 [cs.CL].

Joulin, Armand, Edouard Grave, Piotr Bojanowski and Tomás Mikolov. 2016. “Bag of Tricks for Efficient Text Classification”. *CoRR* abs/1607.01759. arXiv: 1607.01759.

Jungherr, Andreas. 2015. “Twitter use in election campaigns: A systematic literature review”. *Journal of Information Technology & Politics* 13, no. 1 (December): 72–91.

Jungherr, Andreas, Pascal Jürgens and Harald Schoen. 2012. “Why the Pirate Party Won the German Election of 2009 or The Trouble With Predictions: A Response to Tumasjan, A., Sprenger, T. O., Sander, P. G., & Welpe, I. M. “Predicting Elections With Twitter: What 140 Characters Reveal About Political Sentiment””. *Social Science Computer Review* 30 (2): 229–234.

Jungherr, Andreas and Yannis Theocharis. 2017. “The empiricist’s challenge: Asking meaningful questions in political science in the age of big data”. *Journal of Information Technology & Politics* 14, no. 2 (April): 97–109.

Kitschelt, Herbert. 1989. “The Internal Politics of Parties: The Law of Curvilinear Disparity Revisited”. *Political Studies* 37, no. 3 (September): 400–421.

Klinger, Ulrike, Karolina Koc-Michalska and Uta Russmann. 2022. “Are Campaigns Getting Uglier, and Who Is to Blame? Negativity, Dramatization and Populism on Facebook in the 2014 and 2019 EP Election Campaigns”. *Political Communication* (October): 1–20.

Klinger, Ulrike and Uta Russmann. 2017. “Beer is more efficient than social media—Political parties and strategic communication in Austrian and Swiss national elections”. *Journal of Information Technology & Politics* 14, no. 4 (October): 299–313.

Klofstad, Casey A., Anand Edward Sokhey and Scott D. McClurg. 2013. “Disagreeing about Disagreement: How Conflict in Social Networks Affects Political Behavior”. *American Journal of Political Science* 57 (1): 120–134.

Kobayashi, Tetsuro and Yu Ichifuji. 2015. “Tweets That Matter: Evidence From a Randomized Field Experiment in Japan”. *Political Communication* 32, no. 4 (June): 574–593.

Koc-Michalska, Karolina, Darren G. Lilleker, Pawel Surowiec and Pawel Baranowski. 2014. “Poland’s 2011 Online Election Campaign: New Tools, New Professionalism, New Ways to Win Votes”. *Journal of Information Technology & Politics* 11, no. 2 (April): 186–205.

Kreiss, Daniel, Regina G. Lawrence and Shannon C. McGregor. 2017. “In Their Own Words: Political Practitioner Accounts of Candidates, Audiences, Affordances, Genres, and Timing in Strategic Social Media Use”. *Political Communication* 35, no. 1 (October): 8–31.

Kriesi, Hanspeter and Alexander H. Trechsel. 2008. *The Politics of Switzerland.* Cambridge University Press.

Larsson, Anders Olof and Hallvard Moe. 2012. “Studying political microblogging: Twitter users in the 2010 Swedish election campaign”. *New Media & Society* 14, no. 5 (August): 729–747.

———. 2013. “Representation or Participation? Twitter Use During the 2011 Danish Election Campaign”. *Javnost - The Public* 20, no. 1 (January): 71–88.

Lauderdale, Benjamin E. and Alexander Herzog. 2016. “Measuring Political Positions from Legislative Speech”. *Political Analysis* 24 (3): 374–394.

Lee, Eun-Ju. 2013. “Effectiveness of Politicians’ Soft Campaign on Twitter Versus TV: Cognitive and Experiential Routes”. *Journal of Communication* (August): 953–974.

Lee, Eun-Ju and Soo Youn Oh. 2013. “Seek and You Shall Find? How Need for Orientation Moderates Knowledge Gain from Twitter Use”. *Journal of Communication* 63, no. 4 (June): 745–765.

Lehmann, Pola and Malisa Zobel. 2018. "Positions and saliency of immigration in party manifestos: A novel dataset using crowd coding". *European Journal of Political Research* 57 (4): 1056–1083.

Leifeld, Philip. 2013. "texreg: Conversion of Statistical Model Output in R to LaTeX and HTML Tables". *Journal of Statistical Software* 55 (8): 1–24.

Lindstädt, René, Jonathan B. Slapin and Ryan J. Vander Wielen. 2011. "Balancing Competing Demands: Position Taking and Election Proximity in the European Parliament". *Legislative Studies Quarterly* 36, no. 1 (February): 37–70.

Loshchilov, Ilya and Frank Hutter. 2017. "Decoupled Weight Decay Regularization", arXiv: 1711.05101 [cs.LG].

Lucas, Christopher, Richard A. Nielsen, Margaret E. Roberts, Brandon M. Stewart, Alex Storer and Dustin Tingley. 2015. "Computer-Assisted Text Analysis for Comparative Politics". *Political Analysis* 23 (2): 254–277.

Mair, Peter. 2008. "The Challenge to Party Government". *West European Politics* 31, nos. 1-2 (January): 211–234.

Marin, Alexandra and Barry Wellman. 2011. "Social Network Analysis: An Introduction". Chap. 2 in *The SAGE Handbook of Social Network Analysis,* edited by Peter J. Carrington and John Scott, 11–25. London: SAGE.

Martin, Gregory J. and Steven W. Webster. 2018. "Does residential sorting explain geographic polarization?" *Political Science Research and Methods* 8, no. 2 (October): 215–231.

Mikolov, Tomas, Ilya Sutskever, Kai Chen, Greg Corrado and Jeffrey Dean. 2013. "Distributed Representations of Words and Phrases and their Compositionality". In *Advances in Neural Information Processing Systems 26,* edited by C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani and K. Q. Weinberger, 3111–3119. http://papers.nips.cc/paper/5021-distributed-representations-of-words-and-phrases-and-their-compositionality.

Moon, David S. and Øivind Bratberg. 2010. "Conceptualising the Multi-Level Party: Two Complementary Approaches". *Politics* 30, no. 1 (February): 52–60.

Morey, Alyssa C., William P. Eveland and Myiah J. Hutchens. 2012. “The “Who” Matters: Types of Interpersonal Relationships and Avoidance of Political Disagreement”. *Political Communication* 29, no. 1 (January): 86–103.

Müller, Jochen. 2013. “On a Short Leash? Sub-National Party Positions between Regional Context and National Party Unity”. *Journal of Elections, Public Opinion & Parties* 23, no. 2 (May): 177–199.

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

Nall, Clayton, Benjamin Schneer and Daniel Carpenter. 2017. “Paths of Recruitment: Rational Social Prospecting in Petition Canvassing”. *American Journal of Political Science* 62, no. 1 (April): 192–209.

NET-Metrix. 2015. *Internetnutzungsverhalten der User in der Schweiz.* Zurich.

Nguyen, Dat Quoc, Thanh Vu and Anh Tuan Nguyen. 2020. “BERTweet: A pre-trained language model for English Tweets”. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations.* Association for Computational Linguistics.

Nickerson, David W. 2008. “Is Voting Contagious? Evidence from Two Field Experiments”. *American Political Science Review* 102, no. 1 (February): 49–57.

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, no. 1 (March): 85–97.

Norris, Pippa. 2000. *A Virtuous Circle: Political Communications in Postindustrial Societies.* Cambridge University Press.

Nulty, Paul, Yannis Theocharis, Sebastian Adrian Popa, Olivier Parnet and Kenneth Benoit. 2016. “Social Media and Political Communication in the 2014 Elections to the European Parliament”. *Electoral Studies* 44 (December): 429–444.

Obholzer, Lukas and William T. Daniel. 2016. "An online electoral connection? How electoral systems condition representatives' social media use". *European Union Politics* 17, no. 3 (July): 387–407.

Pennacchiotti, Marco and Ana-Maria Popescu. 2011. "A Machine Learning Approach to Twitter User Classification". In *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media,* 281–288.

Pennington, Jeffrey, Richard Socher and Christopher D. Manning. 2014. "GloVe: Global Vectors for Word Representation". In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP),* 1532–1543. October. http://www.aclweb.org/anthology/D14-1162.

Peterson, Andrew and Arthur Spirling. 2018. "Classification Accuracy as a Substantive Quantity of Interest: Measuring Polarization in Westminster Systems". *Political Analysis* 26, no. 01 (January): 120–128.

Pontusson, Jonas and Damian Raess. 2012. "How (and Why) Is This Time Different? The Politics of Economic Crisis in Western Europe and the United States". *Annual Review of Political Science* 15, no. 1 (June): 13–33.

Poole, Keith T. and Howard Rosenthal. 1985. "A Spatial Mode for Legislative Roll Call Analysis". *American Journal of Political Science* 29 (2): 357–384.

Reif, Karlheinz and Hermann Schmitt. 1980. "Nine Second-order National Elections: A Conceptual Framework for the Analysis of European Election Results". *European Journal of Political Research* 8, no. 1 (March): 3–44.

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

Sandberg, Linn A. C. and Patrik Öhberg. 2017. "The role of gender in online campaigning: Swedish candidates' motives and use of social media during the European election 2014". *Journal of Information Technology & Politics* 14, no. 4 (October): 314–333.

Scott, John. 2017. *Social Network Analysis.* 4th ed. London: SAGE.

Selb, Peter, Hanspeter Kriesi, Regula Hänggli and Mirko Marr. 2009. "Partisan choices in a direct-democratic campaign". *European Political Science Review* 1, no. 1 (March): 155–172.

Shaw, Daron R. 1999. "A Study of Presidential Campaign Event Effects from 1952 to 1992". *The Journal of Politics* 61, no. 2 (May): 387–422.

Shugart, Matthew Søberg. 2005. "Comparative Electoral Systems Research: The Maturation of a Field and New Challenges Ahead". In *The Politics of Electoral Systems,* edited by Michael Gallagher and Paul Mitchell, 25–56. Oxford: Oxford University Press.

Sinclair, Betsy, Margaret McConnell and Donald P. Green. 2012. "Detecting Spillover Effects: Design and Analysis of Multilevel Experiments". *American Journal of Political Science* 56, no. 4 (May): 1055–1069.

Strom, Kaare. 1990. "A Behavioral Theory of Competitive Political Parties". *American Journal of Political Science* 34, no. 2 (May): 565.

Strömbäck, Jesper. 2008. "Four Phases of Mediatization: An Analysis of the Mediatization of Politics". *The International Journal of Press/Politics* 13, no. 3 (July): 228–246.

Stromer-Galley, Jennifer. 2014. *Presidential Campaigning in the Internet Age.* New York, NY: Oxford University Press.

Sudulich, Maria Laura, Matthew Wall and David M. Farrell. 2013. "Why bother campaigning? Campaign effectiveness in the 2009 European Parliament elections". *Electoral Studies* 32, no. 4 (December): 768–778.

Taagepera, Rein and Matthew Søberg Shugart. 1989. *Seats and Votes: The Effects and Determinants of Electoral Systems.* New Haven, CT: Yale University Press.

Tavits, Margit. 2009. "The Making of Mavericks: Local Loyalties and Party Defection". *Comparative Political Studies* 42, no. 6 (June): 793–815.

Theocharis, Yannis, Pablo Barberá, Zoltán Fazekas, Sebastian Adrian Popa and Olivier Parnet. 2016. "A Bad Workman Blames His Tweets: The Consequences of Citizens' Uncivil Twitter Use When Interacting With Party Candidates". *Journal of Communication* 66, no. 6 (October): 1007–1031.

Thesen, Gunnar. 2014. “Political Agenda Setting as Mediatized Politics? Media–Politics Interactions from a Party and Issue Competition Perspective”. *The International Journal of Press/Politics* 19 (2): 181–201.

Thorlakson, Lori. 2009. “Patterns of Party Integration, Influence and Autonomy in Seven Federations”. *Party Politics* 15, no. 2 (March): 157–177.

Townsley, Joshua. 2018. “Is it worth door-knocking? Evidence from a United Kingdom-based Get Out The Vote (GOTV) field experiment on the effect of party leaflets and canvass visits on voter turnout”. *Political Science Research and Methods* X:1–15.

Traber, Denise, Simon Hug and Pascal Sciarini. 2014. “Party Unity in the Swiss Parliament: The Electoral Connection”. *The Journal of Legislative Studies* 20 (2): 193–215.

Van Houten, Pieter. 2009. “Multi-Level Relations in Political Parties: A Delegation Approach”. *Party Politics* 15, no. 2 (March): 137–156.

Verge, Tània and Raúl Gómez. 2012. “Factionalism in multi-level contexts: When party organization becomes a device”. *Party Politics* 18 (5): 667–685.

Vergeer, Maurice. 2015. “Twitter and Political Campaigning”. *Sociology Compass* 9, no. 9 (September): 745–760.

Vergeer, Maurice and Liesbeth Hermans. 2013. “Campaigning on Twitter: Microblogging and Online Social Networking as Campaign Tools in the 2010 General Elections in the Netherlands”. *Journal of Computer-Mediated Communication* 18, no. 4 (June): 399–419.

Vergeer, Maurice, Liesbeth Hermans and Steven Sams. 2013. “Online social networks and micro-blogging in political campaigning: The exploration of a new campaign tool and a new campaign style”. *Party Politics* 19, no. 3 (May): 477–501.

Volden, Craig and Elizabeth Bergman. 2006. “How Strong Should Our Party Be? Party Member Preferences Over Party Cohesion”. *Legislative Studies Quarterly* 31, no. 1 (February): 71–104.

Wallsten, Kevin. 2010. “‘Yes We Can’: How Online Viewership, Blog Discussion, Campaign Statements, and Mainstream Media Coverage Produced a Viral Video Phenomenon”. *Journal of Information Technology & Politics* 7, nos. 2-3 (May): 163–181.

Wasserman, Stanley and Katherine Faust. 1997. *Social Network Analysis: Methods and Applications.* New York, NY: Cambridge University Press.

Whiteley, Paul F. and Patrick Seyd. 1994. "Local Party Campaigning and Electoral Mobilization in Britain". *The Journal of Politics* 56, no. 1 (February): 242–252.

Wickham, Hadley. 2016. *ggplot2: Elegant Graphics for Data Analysis.* New York: Springer.

Wueest, Bruno and Christian Mueller. 2014. *Tweetocracy Switzerland: Exploring the representativeness, structuration and content of Swiss party politics on Twitter.* Paper presented at the international conference "Smart city: new media, social participation and urban governance".

Yan, Xiaohui, Jiafeng Guo, Yanyan Lan and Xueqi Cheng. 2013. "A biterm topic model for short texts". In *Proceedings of the 22nd international conference on World Wide Web,* 1445–1456.

Zaller, John R. 1992. *The Nature and Origins of Mass Opinion.* Cambridge: Cambridge University Press.

Zhang, Xinyang, Yury Malkov, Omar Florez, Serim Park, Brian McWilliams, Jiawei Han and Ahmed El-Kishky. 2022. "TwHIN-BERT: A Socially-Enriched Pre-trained Language Model for Multilingual Tweet Representations at Twitter" (15 September 2022). arXiv: 2209.07562.

Chapter X

Appendix

A Chapter II: General Appendix

A.1 Hypotheses

We pre-registered the following hypotheses:

- **H1 - one-sided hypothesis: Neighbourhood effects hypothesis** : *Subjects living in non-experimental households in treated streets are more likely to turn out than subjects living in non-experimental households in control streets.*
- **H2: Partisan mobilisation hypothesis** : *The more politically heterogeneous the street, the weaker the spillover effects between experimental and non-experimental households.*
- **H3: Partisan competition hypothesis** : *The more politically heterogeneous the street, the stronger the spillover effects between experimental and non-experimental households.*
- **H4 - one-sided hypothesis: Neighbourhood effects hypothesis** : *The closer the distance between non-experimental households and experimental households in treated streets, the more likely subjects living in non-experimental households are to turn out compared to non-experimental households living at the same distance to experimental-households in control streets.*
- **H5: Partisan mobilisation hypothesis** : *The closer the distance between non-experimental households and experimental households of the same partisan identity in treated streets, the more likely subjects living in non-experimental households are to turn out compared to non-experimental households living at the same distance to experimental-households in control streets.*
- **H6: Partisan competition hypothesis** : *The closer the distance between non-experimental households and experimental households of a rival partisan identity in treated streets, the more likely subjects living in non-experimental households are to turn out compared to non-experimental households living at the same distance to experimental-households in control streets.*

A.2 Background on the Research Site

The partner for this project was the Labour party in a parliamentary constituency in a small city located the southern part of England. We worked with its Member of Parliament and campaign team. Even though the target elections for this experiment are for the larger European Union constituency of South West of England with its seven MEPs, and a local district (city-based) local government election, the Westminster constituency party is responsible for organising campaigning.

Although there has been a MP for the city since 1295, the current boundaries were last changed in 2010. The constituency has 74,955 registered electors. The constituency covers most of the urban area bar two electoral wards.

In the 2010 General Election, which was the one prior to the research taking place in 2014, Labour won the seat with 38.0 per cent of the vote, the Conservative Party came second with 33.0 per cent, Liberal Democrats came third with 20.0 per cent, with the rest of the vote share going to small parties, UKIP, BNP, Green and the Liberal Party. There are thirteen wards within the constituency, nine of which were selected for the experiment (the most Conservative supporting were excluded).

The local authority boundaries extend beyond the parliamentary constituency. It has 62 seats, which are elected in thirds each year, hence the need for a campaign in 2014. In May 2012, the Labour Party became the majority party on the local council.

A.3 Link between inter- and intra- household mobilization

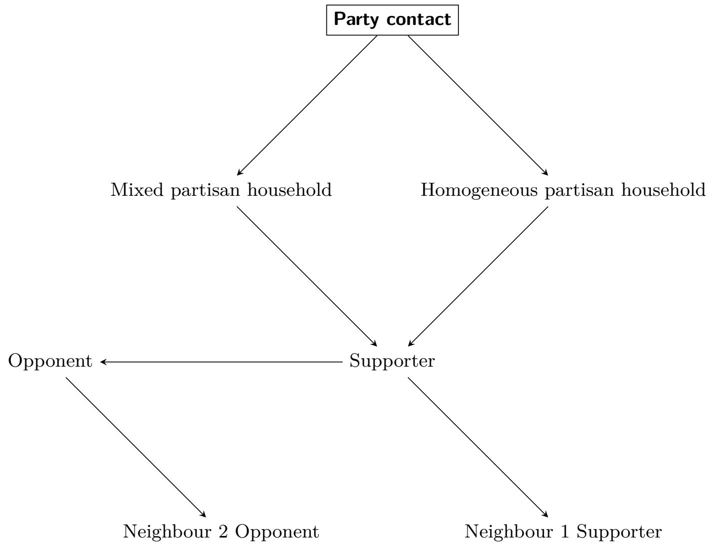
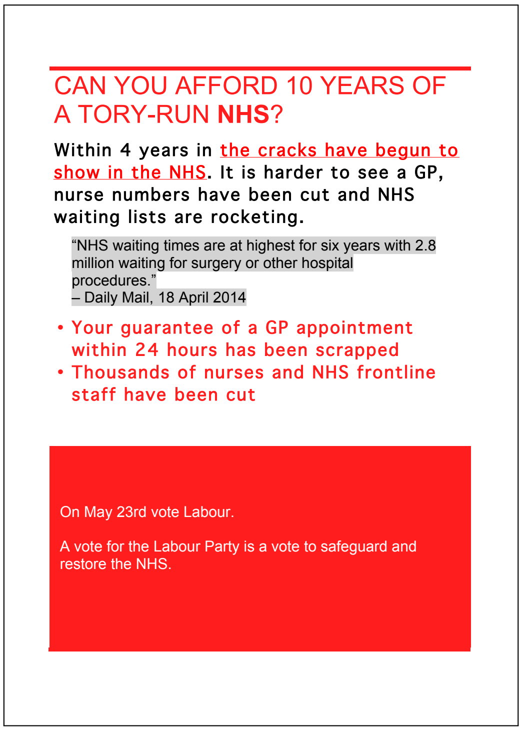


Figure 10: Link between inter- and intra- household mobilization.

A.4 Treatment materials



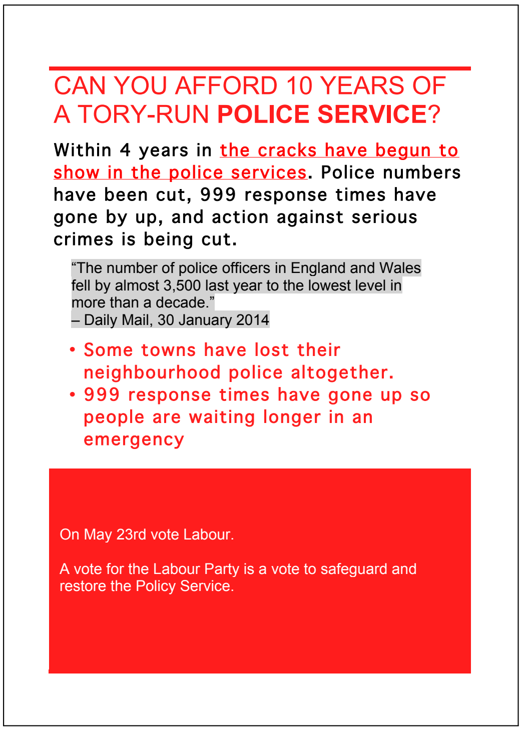


Figure 11: a) Healthcare-themed GOTV leaflet b) Crime- and policing-themed GOTV leaflet.

A.5 Distribution of party self-identifiers

Table 15: Share of party identifiers in experimental and non-experimental samples.

party id	direct	direct_prop	indirect	indirect_prop	total
conservative	628	0.07	896	0.06	1,524
labour	3,381	0.40	5,606	0.35	8,987
nonvoter	196	0.02	2,680	0.17	2,876
other	1,999	0.24	3,103	0.19	5,102
rivalparty	2,171	0.26	3,729	0.23	5,900
total	8,375	1.00	16,014	1.00	24,389

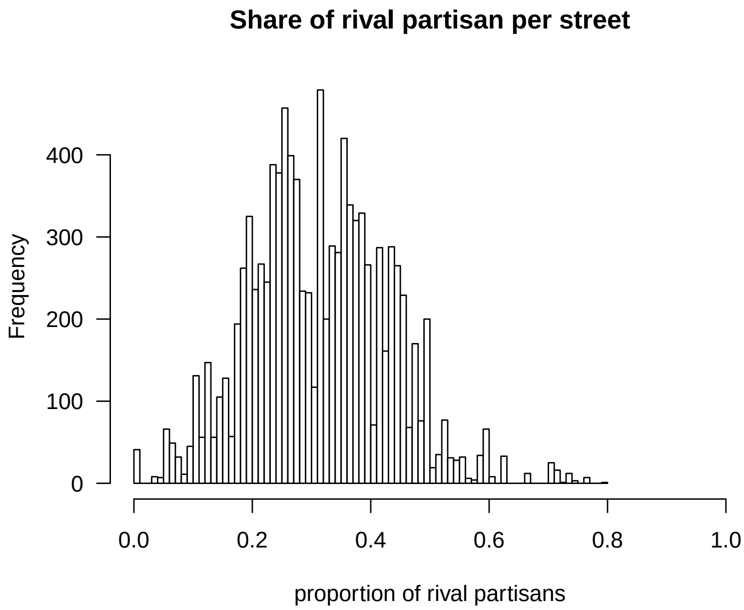


Figure 12: Distribution of neighbours who support a rival party.

A.6 Correlation between share of rival partisans on the same street and share of rival partisans in the most proximate household

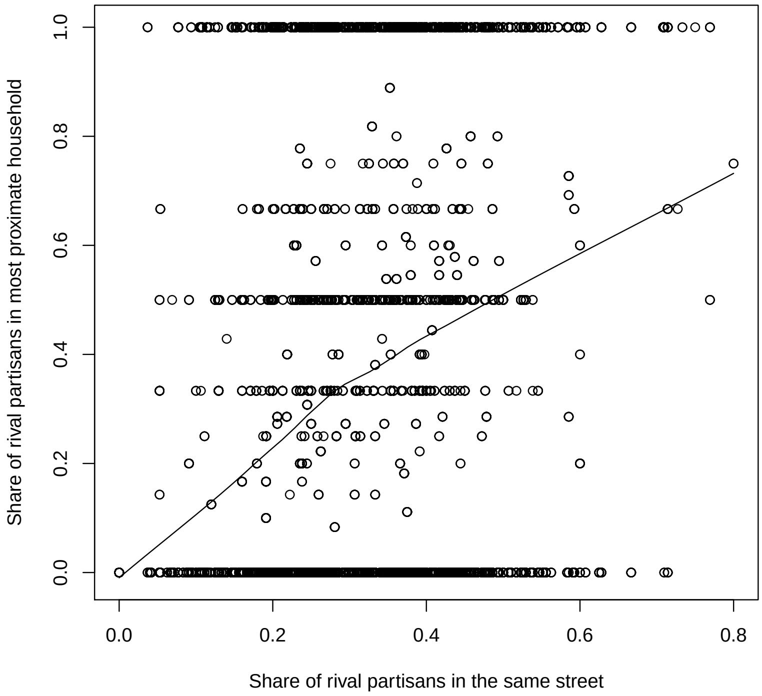
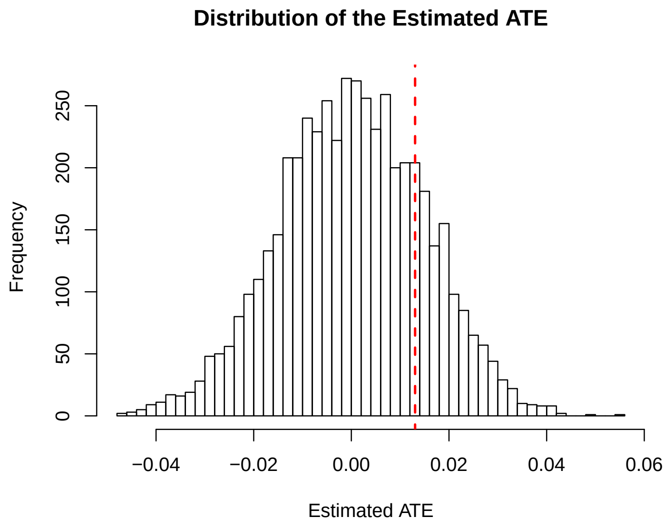


Figure 13: Correlation between share of rival partisans on the same street and share of rival partisans in the most proximate household (kernel smooth function).

A.7 Attrition Check

To check whether individuals in treatment streets are more likely to attrite than individuals in control streets, we estimate the f-statistic from regressing missingness in the outcome variable on assignment to treatment or control streets. We then simulate assignment to treatment and control 5,000 times under the sharp null hypothesis and compare the mean of the f-statistics we obtain under the sharp null to the actual f-statistic from our random assignment.



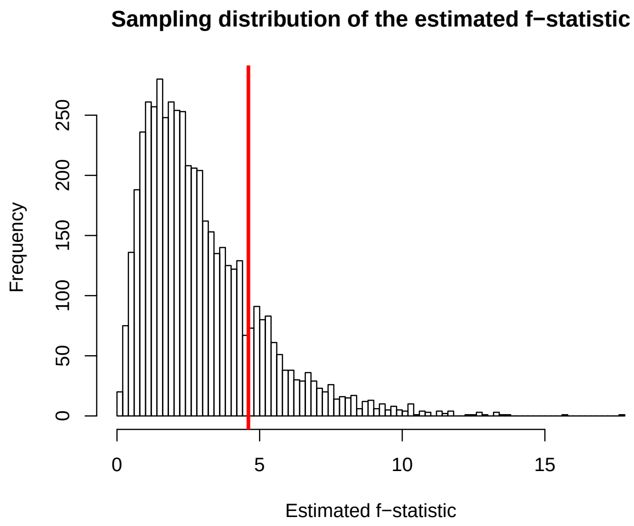
a) Ri-test for differential attrition (p-value: 0.39)

Figure 14: Attrition figure.

A.8 Balance check

Table 16: Balance table.

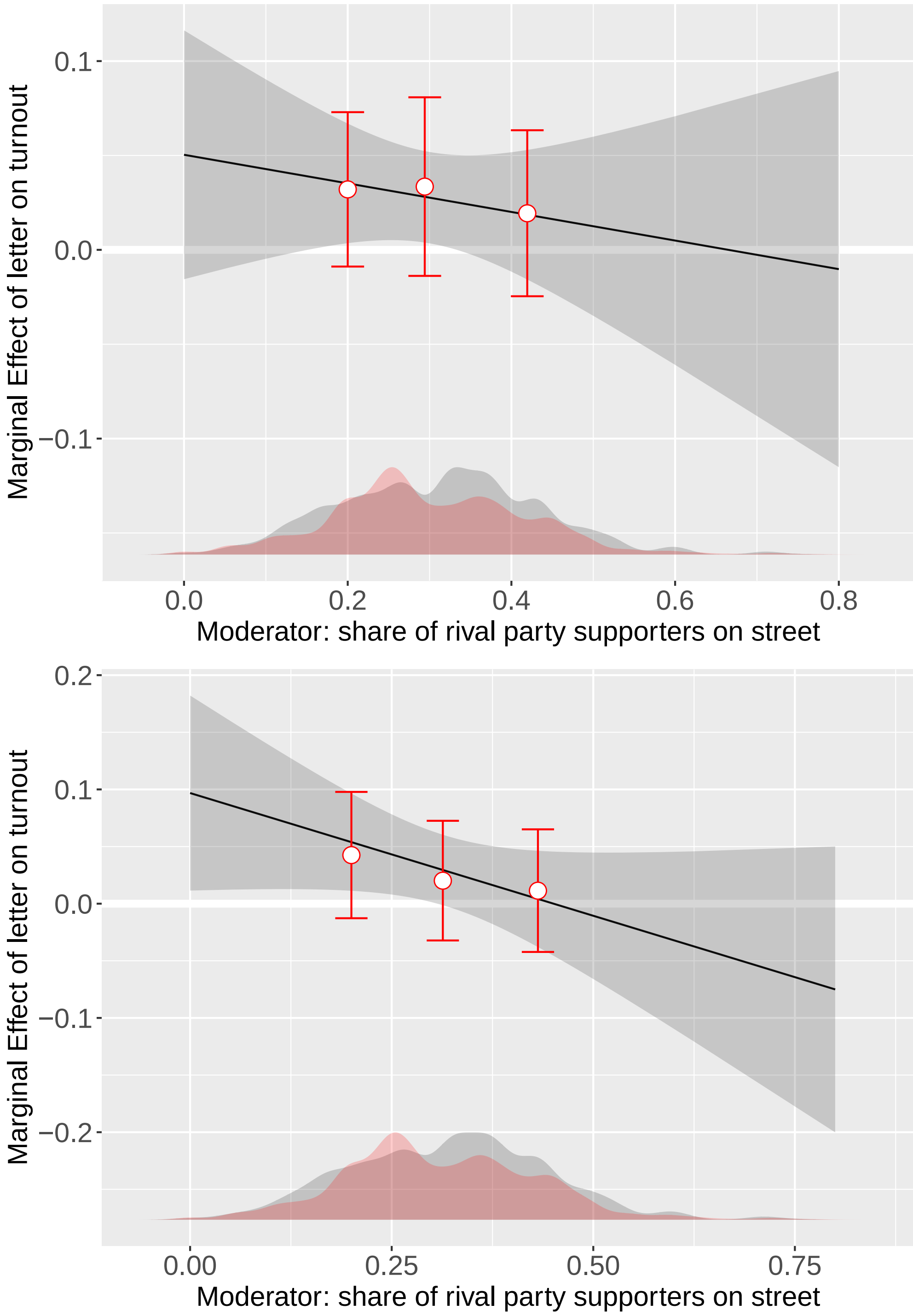
	Control streets	Treatment streets
Household size	2.26	2.22
Male	47.4%	46.3%
Voted 2013	35.6%	36.7%
Labour id	52.2%	55.8%
Rival Party id	38.7%	35.6%
Conservative id	9.1%	8.6%



Ri-test for imbalance on pre-treatment covariates (p-value: 0.18)

Figure 15: Balance figure.

A.9 Robustness checks

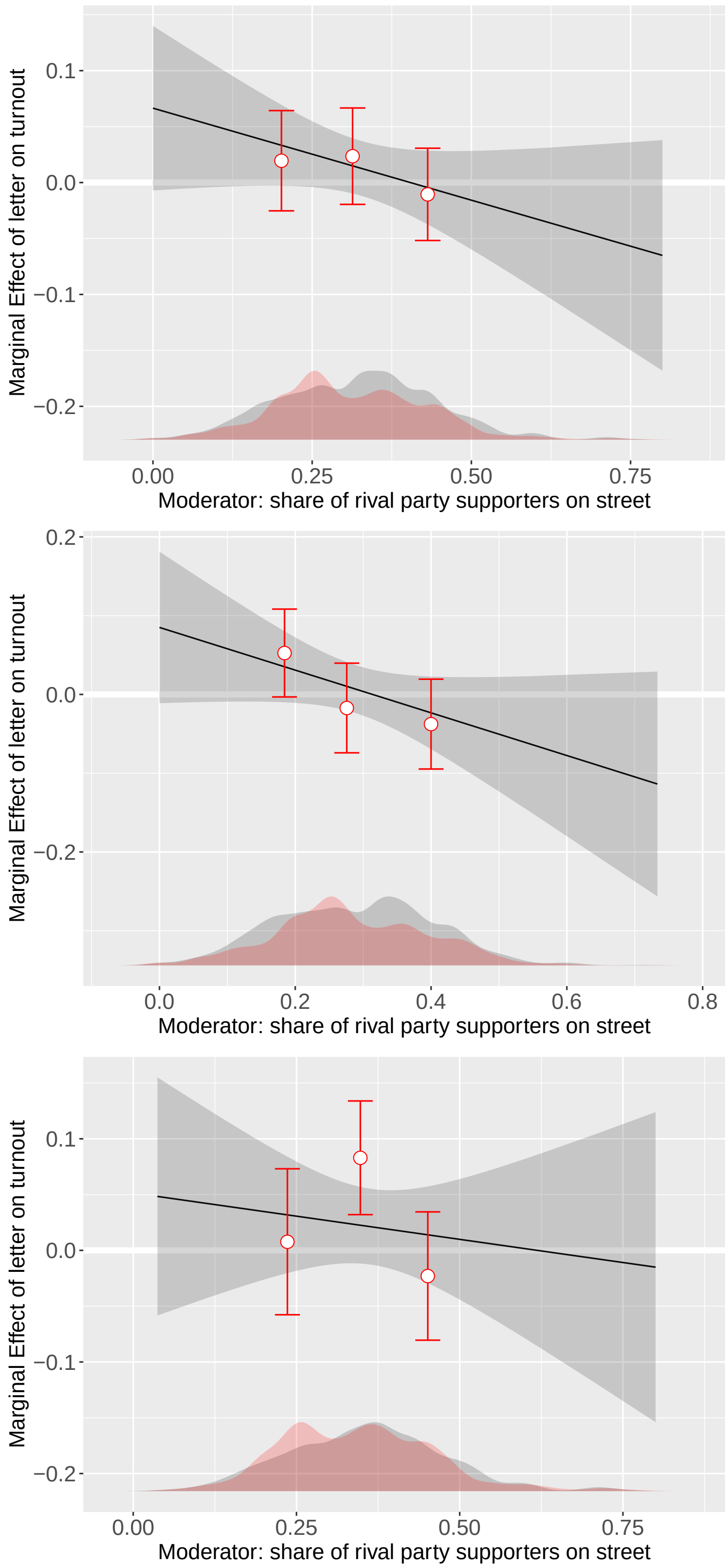


a) including non-identifiers (top) b) all party supporters (bottom)

Figure 16: Marginal effects of leaflet conditional on partisan composition of the neighbourhood - all subjects.

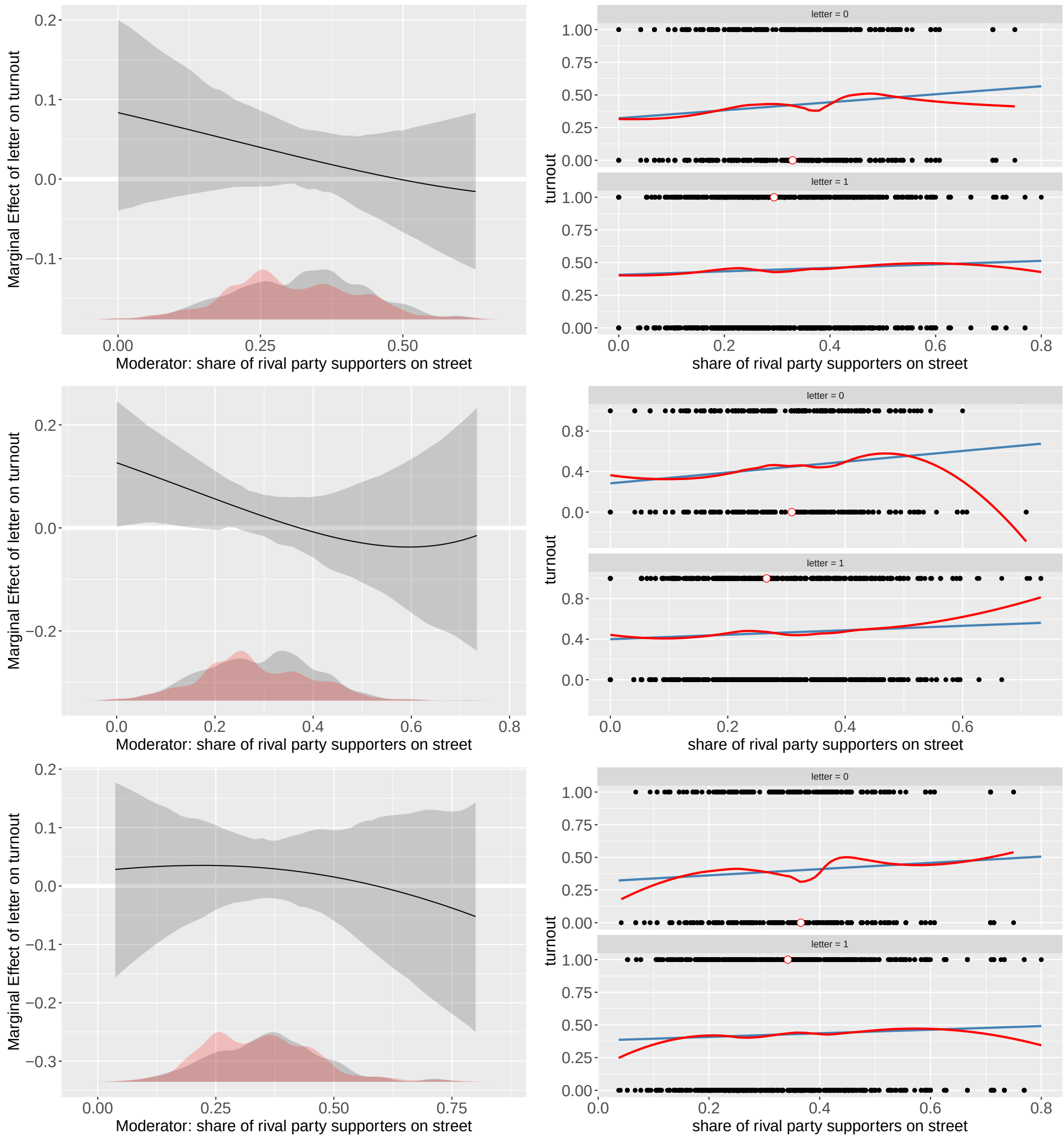
All party supporters

Figure 17: Marginal effects of leaflet conditional on partisan composition of the closest household.



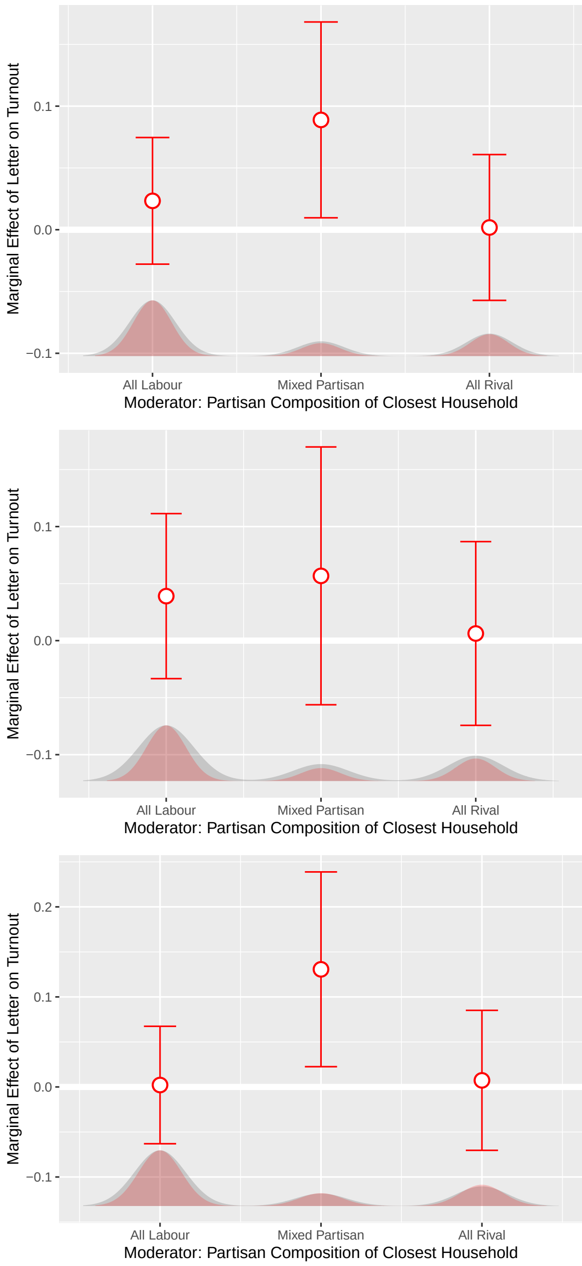
a) all partisans (top) b) Labour supporters (middle row) c) Rival party supporters (bottom)

Figure 18: Marginal effects of leaflet conditional on partisan composition of the neighbourhood—covariate-adjusted.



a) all partisans (top) b) Labour supporters (middle row) c) Rival party supporters (bottom)

Figure 19: Marginal effect of leaflet conditional on partisan composition of the neighbourhood - smooth estimate of the conditional spillover effects on the left; raw data and lowess fit of outcome vs moderator for treated and control groups on the right.



a) all partisans (top) b) Labour supporters (middle row) c) Rival party supporters (bottom)

Figure 20: Marginal effects of leaflet conditional on partisan composition of closest household - covariate-adjusted.

A.10 Other neighborhood-specific factors including dosage

Table 18: Regression of turnout on treatment assignment and partisanship of closest household.

	All partisans	Labour	Rival
Control turnout	0.388	0.417	0.363
	(0.034)	(0.053)	(0.039)
Leaflet	0.033	0.045	0.020
	(0.033)	(0.048)	(0.043)
Closest household mixed	−0.099	−0.068	−0.150
	(0.031)	(0.046)	(0.045)
Closest household rival	0.012	0.023	−0.014
	(0.026)	(0.039)	(0.036)
Closest household mixed x leaflet	0.107	0.043	0.200
	(0.040)	(0.058)	(0.057)
Closest household rival x leaflet	−0.033	−0.050	0.007
	(0.032)	(0.045)	(0.048)
Num. obs.	10006	5497	4509
N cluster	615	615	615

Table 17: Including other neighborhood-specific factors.

	Model I	Model II	Model III	Model IV	Model V	Model VI
Control mean	0.208	0.172	0.203	0.150	0.228	0.185
	(0.038)	(0.046)	(0.059)	(0.067)	(0.053)	(0.066)
Leaflet	0.009	0.055	0.016	0.073	−0.007	0.027
	(0.043)	(0.052)	(0.061)	(0.071)	(0.060)	(0.076)
prop experimental units	0.088	0.039	0.279	0.189	−0.086	−0.120
	(0.119)	(0.120)	(0.163)	(0.170)	(0.148)	(0.155)
prop experimental units x leaflet	−0.082	−0.029	−0.355	−0.257	0.199	0.223
	(0.153)	(0.152)	(0.207)	(0.211)	(0.193)	(0.195)
prop turnout 2013	0.555	0.510	0.536	0.475	0.556	0.511
	(0.131)	(0.130)	(0.196)	(0.191)	(0.157)	(0.163)
prop turnout 2013 x leaflet	0.105	0.169	0.261	0.337	−0.038	0.015
	(0.149)	(0.148)	(0.215)	(0.210)	(0.182)	(0.193)
prop street rival partisan		0.171		0.274		0.168
		(0.118)		(0.170)		(0.182)
tprop street rival partisan x leaflet		−0.239		−0.341		−0.154
		(0.135)		(0.191)		(0.208)
Block fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Cluster standard errors	Yes	Yes	Yes	Yes	Yes	Yes
Num. obs.	10231	10231	5606	5606	4625	4625
N cluster	615	615	615	615	615	615

B Chapter II: Data Appendix

B.1 Selection of the Spillover Sample

Large numbers of households were excluded from the experiment during the design phase. On the one hand, there were legal considerations, such as individuals on no-contact lists. On the other hand, some households had to be excluded for feasability considerations, i.e. to ensure that the share of households could realistically be targeted within the campaign budget. Those excluded households comprise the spillover sample.

The Labour Party implemented the experiment from our randomization and sample selection rules. Our final selection of restrictions on the experimental sample is a carefully checked combination of the restrictions as we designed them and as they were implemented by Labour. Table 19 contains an overview over the restrictions. It shows whether the restriction was implemented in our design code, in Labour's implementation, and in the final selection for the analysis. Those restrictions are discussed in more detail below.

The first two restrictions, **exclude wards with no work going on** and remove individuals who **moved or died**, are elementary. Labour targeted only a selection of nine wards in the constituency. Data from additional wards where no work was being implemented so they were not part of the experiment are excluded at the start. A very small number of individuals in our data were marked as having moved or died before the experiment was implemented. Those individuals are not part of the experiment in any way. Individuals who fall under either category are completely dropped from the data set at the very beginning because they are not relevant to the present experiment. Therefore, they are not part of the analysis and replication materials.

The three following restrictions split the sample into individuals who were directly targeted in the experiment and individuals who were not directly targeted. Individuals who are on **no contact** lists or were missing phone contact information were not part of the experiment because the original experimental design included a telephone survey as one of the outcomes. Therefore, individuals who could not be contacted for legal or practical reasons were not part of the experiment. In a similar vein, **postal voters** were also excluded from being targeted in the experiment. Finally, individuals with certain pre-treatment **partisan-**

Table 19: Overview over sample restrictions.

Description	Our Code	Labour's File	**Final Selection**	Explanation
Exclude wards with no work	Yes	Yes	**Yes**	Those individuals are dropped early in the final code.
Moved or died	Yes	Yes	**Yes**	Individuals who died or moved are dropped early in the final code.
Do not contact	Yes	Yes	**Yes**	We follow our code 1:1, not Labour's, here.
Postal voters	Yes	No	**Yes**	Labour's file does not explicitly mention to exclude postal voters, but careful checking of the implementation data revealed that those individuals were actually excluded as specified in our code.
Partisanship	Yes	No	**Yes**	Individuals with certain voter identifications were excluded across all wards. Additionally, Conservative voters were excluded in two of the wards.
People who cannot vote	No	Yes	**No**	Individuals who cannot vote are not excluded from the experiment but we create a variable `unable_to_vote` which flags people who do not have voting rights.

ship identifications were systematically excluded from the experiment. This was necessary because limitations in campaign funds. Therefore, individuals with a range of pre-treatment party identifications were systematically excluded such that Labour was able to target voters who they identified as more likely Labour supporters. Table 20 gives an overview over the number of individuals who were excluded from the experiment for each of the three restrictions.

Table 20: Summary of Restrictions and Final Sample at the Individual Level.

	Yes	No
Do not contact?	31,810	13,705
Postal voter?	8,564	36,951
Partisanhip restriction?	19,263	26,252
	Indirect	Direct
Individual sample	37,534	7,981

These three restrictions split the sample into individuals who were part of the experiment, i.e. directly targeted, and other who were only indirectly part of the experiment. While the restrictions are defined and implemented on the individual level, the actual level of treatment distribution is the household. Therefore, certain individuals who would not have been part of the experiment were directly targeted as well because at least one other member in the same households was targeted. They were thus exposed to the leaflet (or the control condition) because at least one household member was part of the experiment. The individual-level variable denoting the experiment sample has, therefore, to be expanded to account for this detail. Table 21 contains the number of individuals in each sample before and after adjusting for households.

Table 21: Experiment Sample before and after Household-Level corrections.

	Direct	Indirect
Individual sample	7,981	37,534
Final sample	9,460	36,055

Lastly, some individuals are marked as having **no voting rights** for the election under study. Those individuals should have been excluded from the experiment and the campaign as a whole as they cannot be persuaded in any way for obvious reasons. However, this was not consistently implemented. Thus, we do not exclude these individuals from the experimental sample but we create an indicator variable. Table 22 contains the number of individuals without voting rights broken down for the two samples.

Table 22: Overview of Individuals Without the Right to Vote in the Sample.

	Sample		
Voting Rights?	Direct	Indirect	Total
Yes	9,387	34,833	44,220
No	73	1,222	1,295

B.2 Geolocation and Distance Computation

We use the Google's Geocoding API to locate addresses in our data on the map, i.e. to find their latitude/longitude.

Geolocating addresses is an inherently inexact procedure with several problems. One is that the names of towns and streets are in no way unique even within a region or country and much less worldwide. Google's solution to this problem is what they call "region biasing". We biased the results towards the UK, such that results from the UK are prefered but results are not restricted to be in the town the experiment ran or the UK. Therefore, we put the data from Google through a number of post-processing steps to make sure that results are in the UK and results are in the town

Quality of the data

The Google Maps API gives two indications of the quality of the geolocated address (see here). One is an indicator for whether only a part of the address could be located. The other gives an indication for how precise the geocode is.

A partially matched address means that the complete address in the input does not exist but a subset of the address parts could be matched. Around 20% of the addresses we geolocated are of this type. This is to be expected as many of the address lines contain more than the street address, especially the name or number of the flat, e.g., "Flat 4, 249 Cumberland Road, ...". Although the match for the address is only partial, that is, only "249 Cumberland Road, ..." is identified, the location of the building is still correct.

The precision of the location is returned in four categories. The best quality matches are precise geocodes for the building. If those are unavailable, Google Maps either interpolates the range between two points (e.g. neighbouring street addresses with precise locations), locates the address at the geometric center of a road, or returns an "approximate" location without further specification.

To deal with those two issues, we use two versions of the geolocated data to compute distance-based variables for the analyses. The first one, which is identified by the `all` keyword in the variable name (e.g. `match_top1_all_dist`), uses all available geolocation data as long as it passes the checks described above. This maximizes the amount of available data and thus the power of the analyses.

Identification of Closest Household(s)

The first step is to compute the distance between addresses from the latitude/longitude from the Google Maps API. The distances we compute are "as the crow flies" or "great-circle distances". The algorithm ("Vincenty" ellipsoid) and reference grid (WGS84) results in distances which are exact +/- 1 metre. As treatment assignment was on the street level, we only compute distances and identify closest households within the same street.

For every household which was not targeted directly in the experiment, we identify the closest household(s) which were part of the experiment. This happens in two different modes. The first mode is when there is no household in the experiment at the same address. This can either be because there is only one household in the building or, if there are more than a single household, no household in the given building was part of the experiment. In this case, the closest household is identified by minimizing the distance between the location of the current address and the location of the all addresses with at least one household in the experiment. The second mode is when there are households in the experiment at the same address as the in indirect household to match. This very common for any building, i.e. a single locatable entity, comprising several separate flats. In this case, all households in the same building which were directly part of the experiment are identified as the closest households. If households are matched within buildings, the distance to the match is equal to zero. In this case, we impute the distance with the closest non-zero distance in our data and we generate a dummy variable for whether the matched household(s) are within the same address (e.g. `match_top1_all_within`).

Both modes often lead to ties which is most commonly the case when households are matched within buildings or the address of closest treated household includes additional treated households. In case of a tie, all closest households are averaged over. We then compute the number of individuals (`..._n_obs`) and the number of households (`..._n_hh`) we average over.

B.3 Data Description

Table 23 contains an overview of the number of units split between treated and control across direct and indirect samples.

Table 23: Number of Individuals, Households, and Streets in each Group.

	Direct		Indirect	
	Treated	Control	Treated	Control
Individuals	6,940	2,520	26,667	9,388
Households	3,447	1,250	13,760	4,944
Streets	447	166	447	165

Table 24 contains an overview over the number of individuals for each pre-treatment voter identification as well as the guide on how we recoded and aggregated these voter identifications into variables in our data.

Table 24: Voter Identification Breakdown and Details of Recoding.

		Recoding Details					
Voter Id.	N	..._labour	..._conservative	..._rivalparty	..._other	..._nonvoter	..._observed
labour	11,930	1	0	0	0	0	1
conservative	2,135	0	1	0	0	0	1
against	6,768	0	0	1	0	0	1
bnp	6	0	0	1	0	0	1
green	177	0	0	1	0	0	1
liberal democrat	460	0	0	1	0	0	1
other party	12	0	0	1	0	0	1
plaid cymru	1	0	0	1	0	0	1
respect/soc lab	2	0	0	1	0	0	1
snp	3	0	0	1	0	0	1
uk independence	301	0	0	1	0	0	1
don't know	6,213	0	0	0	1	0	1
independent	1	0	0	0	1	0	1
won't say	297	0	0	0	1	0	1
non voter	3,375	0	0	0	0	1	1
n/a	13,834	0	0	0	0	0	0

C Chapter IV

Table 25: Core set of Swiss political Twitter users.

Name	*Party*	*Name*	*Party*
Bernhard Guhl	BDP	Regula Ryth	GPS
BDP Schweiz	BDP	Robert Cramer	GPS
Martin Landolt	BDP	Adèle Thorens	GPS
Lorenz Hess	BDP	Francine John	GPS
Rosmarie Quardranti	BDP	Anne Mahrer	GPS
Barbara Schmid Federer	CVP	Christian Van Singer	GPS
Marco Romano	CVP	Maya Graf	GPS
Viola Amherd	CVP	Ricardo Lumengo	SPS
Kathy Riklin	CVP	Cédric Wermuth	SPS
Brigitte Haeberli	CVP	SP Schweiz	SPS
Jacques Neirynck	CVP	Christian Levrat	SPS
Luc Barthassat	CVP	Jacqueline Badran	SPS
CVP PDC PPD PCD	CVP	Bea Heim	SPS
Primin Bischof	CVP	Philipp Hadorn	SPS
Brigitte Häberli	CVP	Roberto Zanetti	SPS
Christophe Darbellay	CVP	Carlo Sommaruga	SPS
Elisabeth Schneider	CVP	Jacqueline Fehr	SPS
Yannick Buttet	CVP	Jean-Francois Steiert	SPS
Paul Andre Roux	CVP	Alain Berset	SPS
Jean-René Fournier	CVP	Evi Allemann	SPS
Christian Lohr	CVP	Yvonne Feri	SPS
Alois Gmür	CVP	Pascale Bruderer	SPS
Ida Glanzmann	CVP	Jean Chr. Schwaab	SPS
Stefan Müller	CVP	Edith Graf-Litscher	SPS
Konrad Graber	CVP	Didier Berberat	SPS
Dominique de Buman	CVP	Matthias Aebischer	SPS
Filippo Lombardi	CVP	Silvia Schenker	SPS
Martin Candinas	CVP	Susanne L. Oberholzer	SPS
Ruth Humbel	CVP	Mathias Reynhard	SPS
Stefan Engler	CVP	Geraldine Savary	SPS
Felix Gutzwiler	FDP	Manuel Tomare	SPS
Hugues Hiltpold	FDP	Cesla Amarelle	SPS
FDP.Die Liberalen	FDP	Paul Rechsteiner	SPS
Christian Wasserfallen	FDP	Marra Ada	SPS
Hans-Peter Portmann	FDP	Simonetta Sommaruga	SPS
Isabelle Moret	FDP	Valérie Piller	SPS
Schilliger Peter	FDP	Roger Nordmann	SPS
Theiler Georges	FDP	Martina Munz	SPS
Christa Markwalder	FDP	Jacques-André Maire	SPS
Ignazio Cassis	FDP	Marina Carobbio	SPS
Ruedi Noser	FDP	Rebecca Ruiz	SPS
Filippo Leutenegger	FDP	Eric Nussbaumer	SPS
Petra Gössi	FDP	Andy Tschümperlin	SPS
Andrea Caroni	FDP	Claude Janiak	SPS

Table 25: (continued)

Name	*Party*	*Name*	*Party*
Pierre-André Monnard	FDP	Nadine Masshardt	SPS
Fathi Derder	FDP	Sylvie Perrin-Aquet	SPS
Daniel Stolz	FDP	Stéphane Rossini	SPS
Markus Hutter	FDP	Barbara Gysi	SPS
Johann Schneider-Ammann	FDP	Chantal Galladé	SPS
Doris Fiala	FDP	Claudia Friedl	SPS
Grünliberale Schweiz	GLP	Maria Bernasconi	SPS
Beat Flach	GLP	Ursula Schneider Schneiter	SPS
Isabelle Chevalley	GLP	SVP Schweiz	SVP
Jürg Grossen	GLP	Natalie Rickli	SVP
Roland Fischer	GLP	Andreas Aebi	SVP
Martin Bäumle	GLP	Grin Jean-Pierre	SVP
Thomas Maier	GLP	Luzi Stamm	SVP
Bastien Girod	GPS	Liliane Maury-Pasquier	SVP
Balthasar Glättli	GPS	Thomas de Courten	SVP
Yvonne Gilli	GPS	Pierre Rusconi	SVP
Antonio Hodgers	GPS	Marianne Binder	SVP
Grüne Schweiz	GPS	Christoph Mörgeli	SVP
Aline Trede	GPS	Thomas Hurter	SVP
Ueli Leuenberger	GPS	Oskar Freysinger	SVP
Peter Haag	GPS	Verena Herzog	SVP
Daniel Vischer	GPS	Ulrich Giezendanner	SVP
Jean-Pierre Graber	SVP	Heinz Brand	SVP
Adrian Amstutz	SVP	Maximilian Reimann	SVP
Werner Hösli	SVP	Céline Amaudruz	SVP
Alfred Heer	SVP	Andrea Geissbühler	SVP
Claudio Miotti	SVP	Thomas Hardegger	SPS
Hansjörg Knecht	SVP	Daniel Jositsch	SPS
Yves Nidegger	SVP	Lukas Reimann	SVP
Florin Schütz	SVP	Toni Brunner	SVP

Table 26: Keyword gazetteer for party recognition.

Regular expression	Party	Regex	party
BDP	BDP	JS	SPS
buergerlich.demokratisch	BDP	jungsozialist	SPS
bourgeois.democratique	BDP	juso	SPS
PBD	BDP	sozialdemokrat	SPS
borghese.democratico	BDP	SP	SPS
christdemokrat	CVP	PS	SPS
christlich.demokrat	CVP	socialiste	SPS
christlichdemokratisch	CVP	socialista	SPS
CVP	CVP	giso	SPS
démocrate. *?chretien	CVP	second@	SPS
PDC	CVP	seconda	SPS
democratico.cristiano	CVP	schweizerische.volkspartei	SVP
PPD	CVP	union.democratique	SVP
FDP	FDP	PDB	SVP
freisinn	FDP	unione.democratica.del.centro	SVP
liberale.partei	FDP	UDC	SVP
liberalen	FDP	politi	general
LPS	FDP	gemeinderat	position
liberaux.radicaux	FDP	kantonsrat	position
parti.liberal	FDP	landamman	position
PLR	FDP	landrat	position
liberali.radicali	FDP	nationalrat	position
popolare.democratico	FDP	regierungspräsident	position
GLP	GLP	regierungsrat	position
gruen.liberal	GLP	schultheiss	position
grünliberal	GLP	staatsrat	position
PVL	GLP	ständerat	position
vert.libéral	GLP	standeskommission	position
vert.liberaux	GLP	conseil.des.états	position
verdi.liberali	GLP	conseil.d.état	position
GP	GPS	conseil.exécutif	position
gruene	GPS	conseiller.aux.éetats	position
grüne	GPS	conseiller.municipal	position
ökoliberal	GPS	conseil.municipal	position
écologiste	GPS	conseil.national	position
parti.ecologiste	GPS	grand.conseil	position
verts	GPS	consiglio.comunale	position
Ecologista	GPS	consiglio.degli.stati	position
		granz.consiglio	position

Table 27: User and network descriptives.

User statistics	Average	Std. Dev.
Tweet count	710	2,043
Follower count	721	2.663
Network statistics		
Indegree	53.9	69.3
Outdegree	53.9	59.9
Shortest path	2.5	0.7

Table 28: Main indicators.

Party	N	Vote share (%)[a]
Social Democratic Party	361	18.8
The Liberals	245	16.4
Christian Democratic People's Party	223	11.6
Swiss People's Party	206	29.4
The Greens	155	7.1
Green-Liberal Party	122	4.6
Conservative Democratic Party	72	4.1
Language		
German	805	
French	420	
English	116	
Italian	43	
Level		
National	201	
Cantonal	552	
Municipal	631	
Gender		
Male	686	
Female	248	
Organization	450	

[a] Vote share in the 2015 elections to the National Council.

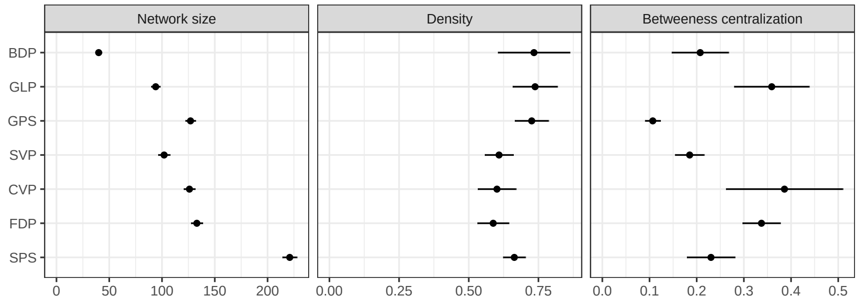


Figure 21: Programmatic cohesion: @-Mentions.

D Chapter VI

Table 29: Source for election lists by country.

Country	Source	Retrieved
Austria	Official publication	9 May 2019
Belgium	Official publication	5 May 2019
Bulgaria	Official publication	9 May 2019
Croatia	Official publication	10 May 2019
Cyprus	Party website(s)	12 May 2019
Czechia	Official publication	6 May 2019
Denmark	Official publication	9 May 2019
Estonia	Official publication	11 May 2019
Finland	Official publication	10 May 2019
France	Party website(s)[a]	4 April 2019
Germany	Party website(s)[a]	30 March–9 April 2019
Greece	Official publication	20 May 2019
Hungary	Official publication	7–11 May 2019
Ireland	Official ballot paper	13 May 2019
Italy	Official list	14 May 2019
Lithuania	Official publication	11 May 2019
Luxembourg	Party website(s)[a]	12 May 2019
Latvia	Official publication	11 May 2019
Malta	Party website(s)[a]	12 May 2019
Netherlands	Party website(s)[a]	3 May 2019
Poland	Official regional ballots	2 May 2019
Portugal	Party website(s)[a]	7–13 May 2019
Romania	Official list	3 May 2019
Slovakia	Official list	10 May 2019
Slovenia	Official list	11 May 2019
Spain	Official list	15 May 2019
Sweden	Official list	9 May 2019
United Kingdom	Official ballot	13 May 2019

[a] List was validated after official publication.

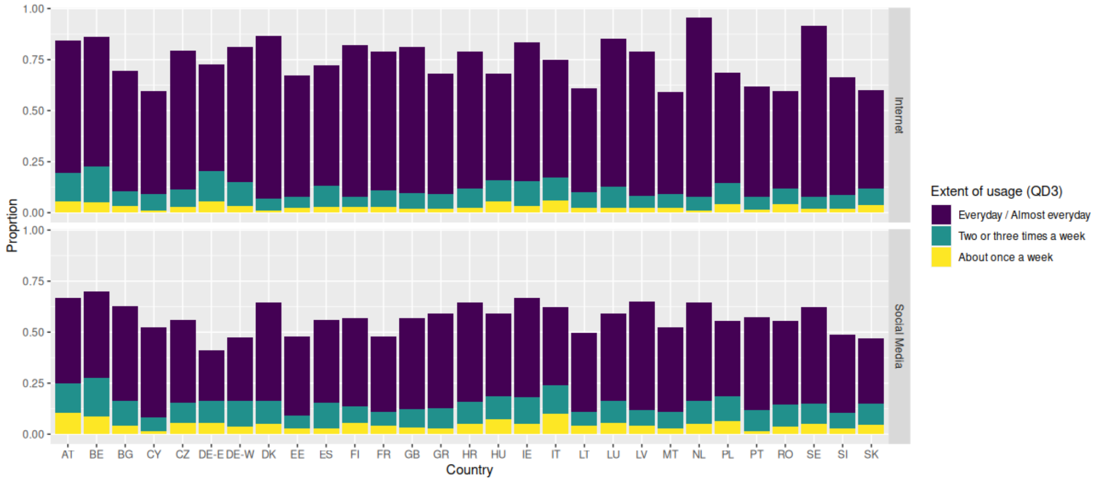


Figure 22: Comparison of internet and social media usage variables in the Eurobarometer survey.

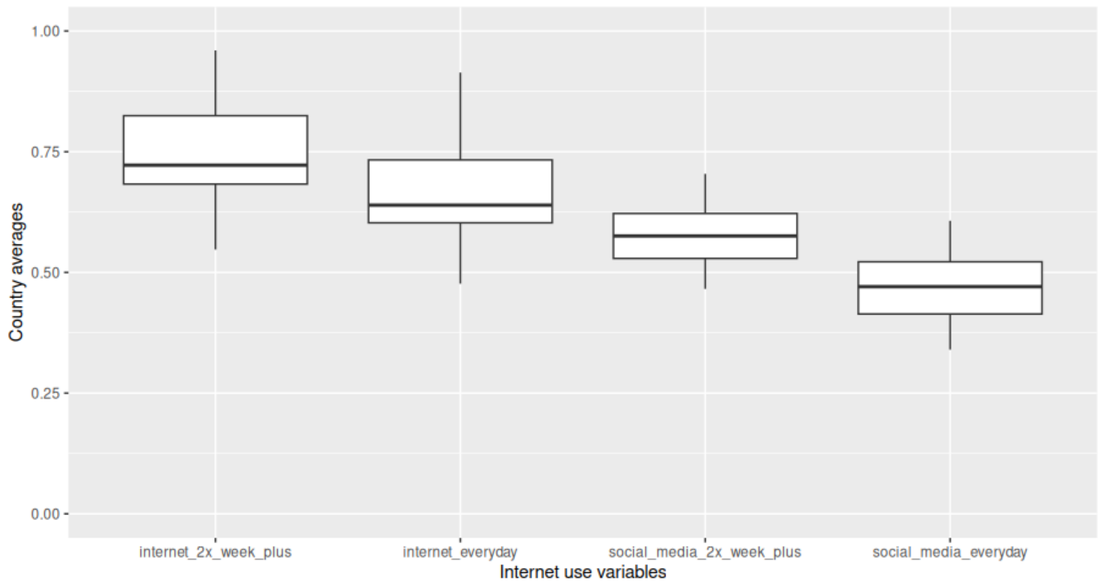


Figure 23: Comparison of internet and social media usage variables in the dataset.

Table 30: Validation data.

		Parties	
Country	Complete	Overall	Complete
Austria		7	2
Belgium	✓	12	12
Bulgaria		7	4
Croatia		6	2
Cyprus	✓	5	5
Czechia		8	5
Denmark	✓	7	7
Estonia		5	2
Finland		7	2
France		8	3
Germany		7	6
Greece	✓	6	6
Hungary		5	2
Ireland	✓	14	14
Italy		4	1
Latvia		9	2
Lithuania		8	1
Luxembourg		6	4
Malta	✓	2	2
Netherlands		10	2
Poland		9	2
Portugal		7	3
Romania		7	2
Slovakia		10	2
Slovenia		7	2
Spain		9	2
Sweden		8	2
United Kingdom	✓	24	24
Total	7	224	123

E Chapter VII

E.1 Tweet Classification Codebook

The following pages show the codebook. A preview of the layout of the layout as it appeared to coders on Qualtrics is below followed by the complete codebook directly from the original HTML file.

Please read the following instructions carefully. It is very important that you understand how to classify Tweets and how we define the categories.

How to classify a Tweet?

Identify whether at least one of the categories described below applies to the Tweet. If more than one applies, select the category which is more important to the message in the Tweet.

Keep in mind that some Tweets will fall under the **unrelated/other** category (see description below) because they are too short or unspecific, or do fit any other category.

What are the different categories of Tweets?

1. Narrating: Narration of the whereabouts of the candidate
2. External messaging: Sending messages or engaging in conversations
3. Requesting action: Asking the public to do something for the candidate
4. Thanking: Thanking the public for their support
5. Positioning: Statements about the candidate's policy positions
6. Directing to information: Sharing external sources of information
7. Unrelated/other: No relation to the elections or no category applies

---

Narrating: Narration of the whereabouts of the candidate

The narrating category contains Tweets which describe where a candidate is at the time or where he/she plans to be. Often these are events where the candidate **meets people** in a **specific place** and/or **time**. Sometimes, these can be bigger events with several candidates, such as a debate.

Example 1

> I'll be in Greystones tomorrow morning, looking forward to hearing what European issues matter to commuters. #doyleforeurope #EUelections2019

Please read the following instructions carefully. It is very important that you understand how to classify Tweets and how we define the categories.

How to classify a Tweet?

Identify whether at least one of the categories described below applies to the Tweet. If more than one applies, select the category which is more important to the message in the Tweet.

Keep in mind that some Tweets will fall under the **unrelated/other** category (see description below) because they are too short or unspecific, or do fit any other category.

What are the different categories of Tweets?

1. Narrating: Narration of the whereabouts of the candidate
2. External messaging: Sending messages or engaging in conversations
3. Requesting action: Asking the public to do something for the candidate
4. Thanking: Thanking the public for their support
5. Positioning: Statements about the candidate's policy positions
6. Directing to information: Sharing external sources of information
7. Unrelated/other: No relation to the elections or no category applies

Narrating: Narration of the whereabouts of the candidate

The narrating category contains Tweets which describe where a candidate is at the time or where he/she plans to be. Often these are events where the candidate **meets people** in a **specific place** and/or **time**. Sometimes, these can be bigger events with several candidates, such as a debate.

Example 1

> I'll be in Greystones tomorrow morning, looking forward to hearing what European issues matter to commuters. #doyleforeurope #EUelections2019

This is a narration Tweet because it references **Greystones**, a specific place, and **tomorrow morning**, a time. It mentions talking about **what European issues matter** with a group of people, **commuters**.

Example 2

> #Grimbsy tomorrow 15th May! Come and meet up with your

> @brexitparty_uk MEP Candidates! Cueworld DN31 1BD, 6pm. See you there!

This is a narration Tweet because it references a place, **Grimbsy**, and invites people to **come and meet up**. In this case, it is not just a single candidate but all candidates from a party who take part in the event. This Tweet also includes an image with the names and pictures of each candidate.

External messaging: Sending messages or engaging in conversations

The external messaging category contains Tweets which are part of a **conversation** between a candidate and other Twitter users. The other user(s) do not need to be addressed directly, such as by name or with a mention on Twitter (`@user123`), but it needs to be clear that the message is directed to certain individuals. There is no specific content the Tweet needs to have. Such Tweets are often part of a longer conversation, in which case the Tweet can be the **first message** initiating a conversation or a **reply** to a previous interaction.

Example 1

> @Mozzarp Aw Mollie!! I miss you and I miss our class x Absolute pleasure and it's great to keep in touch!

This is a external messaging Tweet because it is directed to a specific person (**Mollie**) and the message is what a normal conversation between two people looks like. In this case, the candidate met the other user (**Mollie**) during **class** and plans **to keep in touch**.

Example 2

> @LikesPolitics @NewEuropeans You're a troll CherryB & will be blocked

This is a external messaging Tweet because it is part of a conversation between the candidate and another user (**CherryB**). It is a conversation because there has been some interaction between the candidate and **CherryB** before. The candidate would not refer to them as a **troll** or tell them that they **will be blocked** otherwise.

Requesting action: Asking the public to do something for the candidate

The requesting action category contains Tweets in which a candidate asks the public to do something for them. It can be any type of request, for example asking for help for on a specific task, or general request for support in some form. It can also be a request to vote for the candidate or generally vote in the election.

Example 1

> My message in Poland tonight was a simple one: if you want change, you have to vote in the European elections! #ItsTime

This is a requesting action Tweet because the candidate asks the public to **vote in the European elections**.

Example 2

> We'll be calling voters across the West Midlands today in our phone bank. Come along to the West Mids Labour regional office in West Brom to help us make even more calls! #EUelections2019 #VoteLabour

This is a requesting action Tweet because the candidate asks the public for help with **making even more calls** and to **come along to the West Mids Labour regional office**.

Thanking: Thanking the public for their support

The thanking category contains Tweets in which a candidate thanks the public for something they did for the candidate.

Example 1

> Will I ever grow tired of sharing these funny, uplifting crowdfund comments? Probably not!
>
> A massive thanks to everyone who donated so far - we've raised €4309 in just two days! Our target is €5000, can you chip in?

https://donate.garygannon.ie/

This is a thanking Tweet because the candidate thanks **everyone** who **donated** to their campaign.

Example 2

Or, of course, this incredible group of people who have been incredible campaigners, wonderful friends and my biggest supporters.

Thank you ❤️🌹❤️

This is a thanking Tweet because the candidate thanks their **incredible campaigners**, **wonderful friends** and **biggest supporters** for their help.

Positioning: Statements about the candidate's policy positions

The positioning category contains Tweets in which a candidate makes a statement about an issue they support, goals they want to achieve or plans they have if they get elected. The statements can be ideas or positions they support, general plans of focusing on a topic, or working towards a solution for a group of people.

Example 1

Walking the walk - Proud to sign @AutismEurope 's pledge for #MEP candidates! If elected, creating a truly inclusive society will be one of my main goals 👊

#KburinbPajjiżna | #EP2019 🇲🇹🇪🇺

This is a positioning Tweet because the candidate states that **creating a truly inclusive society** is one their **main goals** if they get elected.

Example 2

As NI enters a new talks process, its essential we deliver certainty to local economy.

The decision of Bombardier to sell its Belfast operations is concerning as it leaves 3,600 jobs at risk.

@uuponline is committed to protecting the livelihoods of those affected by this sale

This is a positioning Tweet because the candidate states a clear position that **its essential we deliver certainty to local economy**. The Tweet continues with a specific example of Bombardier's decision **to sell its Belfast operations** and the commitment of the candidate, or the party **@uuponline** in this case, to **protecting the livelihoods of those affected**.

Directing to information: Sharing external sources of information

The directing to information category contains Tweets in which a candidate points the public to pieces of information. It is usually a link to another website but it can also be a description of where to find a piece of information. The information can be a text that they wrote themselves, such as a blog post, or it can be something else that they want the public to read.

Example 1

'Not since Henry VIII's Reformation Parliament has a parliament been so associated with a single cause as this one, & the next month will determine whether it is remembered for catastrophe or salvation.' My @TheNewEuropean column, out tomorrow

This is a directing to information Tweet because it directs the public to look at a specific piece of information, **My @TheNewEuropean column, out tomorrow**. This Tweet does not contain a link to the piece of information but it clearly describes where to find it.

Example 2

> Patriotic politicians, including Geert Wilders and Marine Le Pen, promise to battle the EU https://www.wsj.com/articles/in-search-of-influence-europes-nationalists-reach-toward-a-pact-11556221397?reflink=share_mobilewebshare

This is a directing to information Tweet because it contains a link to an article published in the Wall Street Journal (**wsj.com**). In this case, the preview shows the article's title, **"With Mainstream Parties Struggling, Europe's Nationalists ..."**.

Unrelated/other: No relation to the elections or no category applies

The unrelated/other category contains Tweets which do not clearly fit into any other category. A number of Tweets will fall into the unrelated/other category because many candidates send Tweets which are not related to the 2019 European Parliament Elections. This category cannot be accurately defined but there are two broad types of Tweets that fall into it. On the one hand, Tweets which are not about the election campaign or politics in general. On the other hand, Tweets which might contain parts of one (or several) categories but do not clearly fall into any single one. Sometimes, a Tweet is too short or missing context to clearly establish a category.

Example 1

> Two nights running football delivers unparalleled entertainment. Congratulations to Spurs. @SebDance @Claude_Moraes

This is a unrelated/other Tweet because it is clearly about sports/football, therefore unrelated to the 2019 European Parliament Elections.

Example 2

> May will resign as party leader on 7th of June. She will remain as PM until a successor is chosen.

This is a unrelated/other Tweet because it is a statement that is about politics but it is not related to the election campaign.

Table 31: Negative Binomial Model for Number of Tweets (Outcome) based on Original Tweets (without retweets and quoted tweets).

	Two-month campaign		One-month campaign	
	All Candidates	Incumbents	All Candidates	Incumbents
(Intercept)	0.49*	0.24	0.12	−0.18
	(0.24)	(0.45)	(0.27)	(0.53)
Internet Usage[a]	0.00	0.00	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Citizens Represented (avg)[b]	0.19	0.24	0.25	0.36*
	(0.12)	(0.12)	(0.15)	(0.15)
Open List	−0.19	−0.17	−0.15	−0.21
	(0.38)	(0.34)	(0.45)	(0.44)
District Size (avg)[b]	−0.31*	−0.21	−0.33*	−0.25
	(0.14)	(0.12)	(0.16)	(0.15)
Open List × District Size (avg)[b]	0.37	−0.11	0.48	−0.41
	(0.66)	(0.64)	(0.79)	(0.80)
In Government	−0.09	−0.11	−0.10	−0.17
	(0.09)	(0.13)	(0.09)	(0.15)
Vote Share	−0.00	0.00	−0.00	−0.00
	(0.00)	(0.01)	(0.00)	(0.01)
Left–Right	0.01	0.03	0.03	0.06
	(0.03)	(0.05)	(0.03)	(0.05)
GAL–TAN	−0.00	−0.06	−0.02	−0.08
	(0.03)	(0.05)	(0.03)	(0.05)
EU Against–Pro	0.00	−0.02	0.01	−0.01
	(0.01)	(0.03)	(0.02)	(0.03)
Gender: Male	0.05	−0.06	0.02	−0.10
	(0.05)	(0.11)	(0.05)	(0.12)
EP Spitzenkandidat	0.43	0.42	0.47	0.56
	(0.39)	(0.39)	(0.42)	(0.44)
Incumbent	−0.11		−0.07	
	(0.07)		(0.08)	
log(Followers)	0.09***	0.05	0.10***	0.02
	(0.02)	(0.05)	(0.02)	(0.06)
log(Tweets)	0.43***	0.51***	0.38***	0.52***
	(0.02)	(0.06)	(0.02)	(0.07)
Random Effects (Variance):				
Party (Intercept)	0.04	0.02	0.05	0.01
Country (Intercept)	0.15	0.07	0.22	0.15
Candidates	1791	321	1791	321
Parties	163	102	163	102
Countries	28	28	28	28
Dispersion Parameter	1.00	1.33	0.89	1.03
AIC	18244.93	3523.36	16401.55	3206.70
BIC	18349.25	3591.25	16505.87	3274.58
Log Likelihood	−9103.46	−1743.68	−8181.78	−1585.35

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; [a] Centred at the overall (grand) mean; [b] Standardised.