Title
Estimating Extremism: New Measures of Extreme Party Preferences and Issue Positions

Author
Jakli, Laura

Publication Date
2020

Peer reviewed|Thesis/dissertation

Estimating Extremism: New Measures of Extreme Party Preferences and Issue Positions

by

Laura Viktoria Jakli

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Political Science

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Jason Wittenberg, Chair
Professor Gabriel Lenz
Professor Laura Stoker
Professor Katerina Linos

Spring 2020

Estimating Extremism: New Measures of Extreme Party Preferences and Issue Positions

Abstract

Estimating Extremism: New Measures of Extreme Party Preferences and Issue Positions

by

Laura Viktoria Jakli

Doctor of Philosophy in Political Science

University of California, Berkeley

Professor Jason Wittenberg, Chair

This dissertation develops three novel approaches to conceptualize and quantify aspects of political extremism. My main chapters, which are three standalone papers, specifically examine voters' movement toward more extreme parties and issue positions. A majority of the dissertation focuses on measuring far right extremism due to its prevalence in contemporary democracies. However, Chapter 4 also measures far left radicalization.

More broadly, the dissertation is motivated by the discrepancy between publicly expressed and privately held extreme political views. This gap is overlooked in the study of democracies, although it has important implications for measuring and understanding the diffusion of extreme beliefs in democratic electorates.

The first chapter (Chapter 2) addresses this gap between public and private preferences directly, through the original concept of contingent extremism. I posit that there are two broad categories of far right extremists. The first, who I call staunch loyalists, will cast far right votes whenever they are given the opportunity to so. The second, who I call contingent extremists, will only cast a vote when they believe enough other citizens are already doing so. Their support is contingent on perceived party popularity. To empirically test this concept, I field survey experiments in Germany (n=1,991), France (n=1,770) and Hungary (n=1,015), and measure respondents' willingness to identify as far right supporters when randomly assigned to more or less 'favorable' polling information about the party. Contingent extremism is captured through the difference in rates of far right identification in these treatments. To examine geographic variation in contingent extremism, I then match respondents to their electoral districts, and find that contingent extremists live in districts where the far right has weak electoral support. I use this to derive a local party performance threshold—roughly one-fifth of vote share—at which they begin to support the party openly.

Next, I turn to measuring the gap between the publicly and privately stated positions of far right parties. I posit that far right parties are incentivized to present a more mainstream

ideological profile in public communications than they do in private ones. Until recently, information about their private campaign appeals was difficult to obtain, so scholars could not address this gap. In the last year, social media data transparency initiatives have made it possible to explore the microtargeted appeals parties use to mobilize voters online. The second chapter (Chapter 3) details why this new data source is so critical to advancing comparativists' knowledge of far right political strategy. I demonstrate a range of computational tools that can be used to parse the content of ads data, including an unsupervised scaling model of document positions, structural topic modeling, and sentiment analysis. Using these methods, I evaluate the content of more than 68,000 political campaign ads across 11 European countries and 79 political parties, and provide novel insight into the private political campaigns of the far right.

Next, I turn to a chapter on measuring polarization, since polarization plays an important role in the radicalization of democratic electorates. This third chapter (Chapter 4) details a supervised machine learning approach to detect polarization in social media discourse following high salience news events. I operationalize polarization as an increase in the share of extreme political discourse within a partisan network. My approach is a significant deviation from the current literature, which relies on social network analysis to model network structures and draw inferences about polarization. I argue that a machine learning approach is a necessary supplement, because it evaluates what people actually say about politics. I find that everyday discourse on social media is moderately diverse, but following major news events, political discourse becomes more extreme and partisan. The chapter demonstrates the potential of a supervised learning approach to better understand who is susceptible to polarization and when. Moreover, it offers a path forward for comparative studies of polarization.

Finally, Chapter 5 discusses the implications of my findings for extremist politics in democracies and offers future avenues for research. I conclude that far right political views are more common in the electorate than most experts would suggest, in part due to the gap between the public and private expression of extreme viewpoints.

to Paár Zoltán

Contents

List of Figures

List of Tables

Acknowledgments

Special thank you to my dissertation committee, Jason Wittenberg, Katerina Linos, Laura Stoker, and Gabe Lenz for all of their thoughtful mentorship from prospectus to completion. Michaela Mattes, Aila Matanock, and Ron Hassner also contributed greatly to my graduate education. During my time at Berkeley, I had strong support and friendship from fellow graduate students, and would like to thank Melissa Carlson and Melanie Phillips in particular.

I also benefited immensely from my time at Stanford's Center on Democracy, Development, and the Rule of Law and the Stanford Program on Democracy and the Internet. I would like to thank Stephen Stedman, Nate Persily, Frank Fukuyama, Larry Diamond, Erik Jensen, Marcel Fafchamps, and Didi Kuo for their mentorship in that duration.

Thank you to my parents, Gabriella and Antal, and to my sister Olivia for endless love and encouragement. A most special thank you to Gabi nagyi, Magdi nagyi, and Josh for being my brightest lights. Nagyon szeretlek benneteket.

Chapter 1

Introduction

Over the last two decades, extremist politics diffused across the globe. Presidents and prime ministers with far right political agendas now serve as the leaders of eight democracies. They advance extreme policy agendas—fortifying borders and erecting new ones, enacting extreme protectionist economic policies, and leveraging national crises to impose restrictions on civil and political rights. Even when the far right does not lead government, it is increasingly influential in party politics. Far right parties are seated in the parliaments of 23 European countries, and many hold more than one-fifth of total vote share.

Despite the ubiquity of extremism in democratic politics, the electoral successes of the far right continue to be hard to predict. Pollsters systematically underestimate the electoral performance of far right parties prior to elections. The motivations of far right voters also remain unclear. Scholars still disagree on why people support extreme parties, and there are few approaches available for analyzing how voters become radicalized.

This dissertation puts forward three novel approaches to measure political extremism. Each chapter was written as a separate paper that tries to create a new way to conceptualize and quantify the movement toward more extreme parties and issue positions. All three chapters are motivated by the discrepancy between publicly expressed and privately held attitudes that makes the politics of extremism particularly hard to capture.

To preview these studies, Chapter 2 uses survey experiments to measure how perceived party popularity affects whether people are willing to support far right parties. I also use a novel approach to measure how local social environments and political conditions affect people's willingness to identify as far right supporters. I leverage respondent-level geospatial data to link survey respondents to their district-level far right vote share, and use this information to derive a local party legitimacy threshold at which people with far right views are activated as public party supporters.

The study of far right politics also requires accurate data about party positions and appeals on these positions. Chapter 3 applies computational text analysis to the content of digital campaign advertisements, and proposes this as a new way to measure the policy positions of far right parties. I argue that online political ads are important data because the appeals far right parties make to voters privately differ from their publicly held policy positions. The main objective of this chapter is to demonstrate the wide-ranging applications of political ads as text data and illustrate how it can inform long-standing debates on extremism in the party politics literature.

The rise in extremist momentum corresponds to a rise in polarization. Chapter 4 proposes a supervised machine learning approach to detect polarization in social media discourse, where polarization is defined as the increased prevalence of extreme political speech in partisan networks. The chapter details a research design which combines social network analysis (SNA) with a machine learning approach to text analysis to examine political communication across the ideological spectrum of social media users. By combining these methods, I reduce an inferential leap made by SNA studies of polarization. While SNA alone provides insights into the structural properties of communications, it overlooks what members of these networks actually say about politics. I use the social media discussion surrounding three major political events to demonstrate how my method can be used to classify political text and measure polarization in mass discourse.

Finally, Chapter 5 discusses the implications of the findings for extremist politics in democracies and offers future avenues for research. I conclude that political scientists need to take the gap between privately held and publicly expressed preferences seriously. Far right political views are more common in the electorate than suggested by most data. My dissertation puts forward multiple approaches by which to gauge this gap and better understand the distribution of political preferences in democratic electorates.

Chapter 2

Contingent Extremism: How Perceptions of Party Popularity Activate Far Right Support

2.1 Abstract

For decades, European far right parties existed at the political fringe, garnering the support of a small group of staunch extremists. In recent years, support for these parties has increased at unexpected rates. What explains their political momentum? I argue that far right parties have broadened their base by mobilizing contingent extremists—supporters who have long held extreme beliefs, but who viewed the party as illegitimate in more hostile political opinion climates. I posit that as the perceived popularity of the far right increases, it activates supporters—especially in places where the actual popularity is low. To test this theory, I field experiments in Germany (n=1,991), France (n=1,770), and Hungary (n=1,015) to measure respondents' willingness to identify as far right supporters when assigned to more or less 'favorable' information about far right party popularity through randomly varied polls. I capture contingent extremism through respondents' differential willingness to identify as far right supporters in these experimental polling treatments. Using geospatial analysis, I also find that most contingent extremists reside in voting districts where the far right is electorally weak. Once the party gains approximately one-fifth of the vote share at the district level, respondents from those districts support the party openly.

2.2 Introduction: The Rise of the Far Right

Over the last several election cycles, European countries have shifted from centrist politics rightward. Averaging across European parliamentary elections, far right parties won just under 5 percent of the overall vote share in 1997, then 7.5 percent in 2007, and almost 15 percent by 2017. Far right parties are now represented in twenty-three European countries' parliaments, and in several of these countries, they capture one-fifth of the electorate's vote. Figure 2.1 illustrates the rise of Europe's far right over the last three decades.

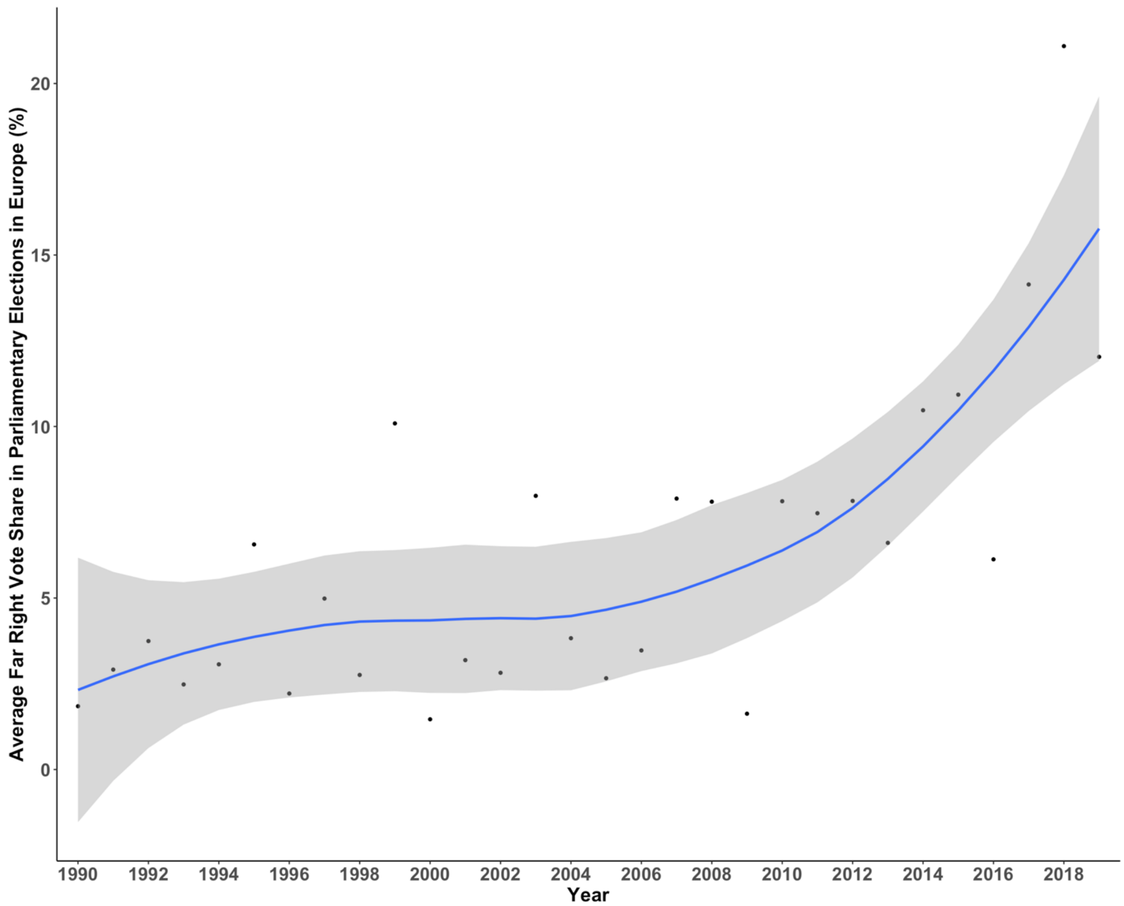


Figure 2.1: The Political Momentum of Europe's Far Right (1990-2019)

Average far right vote share (with LOESS smoothing) in parliamentary elections across 23 European countries.

Although this study examines Europe, the rise of the right is a global phenomenon. In Latin America, almost a dozen far right political parties are active, while Jair Bolsanaro leads Brazil as the country's sitting president (Kernecker and Wagner 2019). In Southeast Asia, Rodrigo Duterte serves as the president of the Philippines, and similarly militant political sentiment has emerged in Thailand, Cambodia, and Myanmar. U.S. President Donald Trump was elected on more extreme policy positions than conservative candidates had voiced in decades, with 27 percent of the voting-eligible U.S. population voting him into office. Far right parties and candidates have rapidly established themselves as political contenders worldwide.

Recent research shows that both the ‘supply’ (e.g., party strategy) and ‘demand’ (e.g., public attitudes) for far right politics have been relatively stable over time (Bonikowski 2017). If far right parties have been stable in their discourse for decades and voters have long held extreme beliefs, why has far right vote share increased rapidly in recent years?

I posit that far right party momentum is due in part to the successful mobilization of contingent extremists. I define contingent extremists as people who have long held extreme beliefs but did not view far right parties as legitimate political contenders in opinion climates hostile to right-wing extremism. I posit that contingent extremists are particularly attentive to public opinion and social norms, and in turn, to the stigma associated with the far right. When opinion climates shift in favor of these parties, contingent extremists no longer perceive the party as at odds with mainstream values. In turn, supply matches demand and they become active party supporters.

To test this theory, I field experiments in Germany (n=1,991), France (n=1,770), and Hungary (n=1,015) to measure respondents’ willingness to identify as party supporters when assigned to more or less ‘favorable’ information about far right party popularity through randomly varied polls. I capture contingent extremism through respondents’ differential willingness to identify as supporters in these experimental treatments. Using geospatial analysis, I find that contingents extremists reside in voting districts where the far right is electorally weak. This suggests that extremists’ direct social environments are critical to their political activation (or lack thereof). However, these experiments also suggest that new information about party popularity can override the stigmatizing effect of local environments.

Moreover, I find that once these parties capture approximately one-fifth of the vote share at the district level, survey respondents from these districts identify as party supporters openly. I use this threshold to develop a framework for understanding how perceptions of party popularity can boost far right support beyond a narrow, non-contingent base.

This paper proceeds as follows: first, I review existing behavioral scholarship on when and why people act on their political preferences, turning the conversation to opinion climates and social tipping points. Using the literature on ‘impersonal influence,’ I posit that contingent extremists use two primary sources of information to evaluate their opinion climate: (1) their immediate social environments, consisting of their personal networks, and (2) their impersonal information networks, including the media. Second, I describe how these opinion climate drivers can either strengthen or subdue public support for right-wing extremism. Third, I discuss my main theory and hypotheses, and outline how opinion climates shape extremists’ mobilization. Next, the paper outlines the design and implementation of the survey experiments, the hypotheses they test, each country’s respondent pool, and the main experimental findings. Finally, I test the core argument that perceived party popularity affects willingness to express support for the far right. Matching survey respondents’ geolocation data to district-level data on far right electoral performance, I find that most contingent extremists reside in places where the far right is electorally weak. I also discuss cross-national differences in treatment effects and conclude with the political implications of these findings.

2.3 Theoretical Framework: How Opinion Climates and Impersonal Influence Shape Political Norms and Behaviors

How Opinion Climates Shape Contingent Extremism

The idea that people sometimes do not articulate or act on their political preferences is prominent in the social movements literature, particularly in Kuran (1989), Kuran (1991), and Kuran (1995) work on revolutionary thresholds. Kuran theorizes that social factors, which vary with circumstances, may cause people to gain 'reputational utility' from hiding their political preferences (Kuran 1991, 30). This leads to a fragile equilibrium in which relatively minor political developments can set a bandwagon in motion to alter the political order.

The unexpected political momentum of the far right represents the breakdown of a tenuous center-seeking political equilibrium, and I argue that information environments are critical to this development. Exposure to information through personal (primarily local) networks that frame the far right as an unpopular fringe movement can reduce support for right-wing extremism, even amongst those with strong far right political preferences. At the same time, exposure to information through impersonal (primarily national) networks matters, too. Impersonal networks may signal incongruent messages and frame far right parties as legitimate political newcomers that voice popular views. The imbalances between these two information sources can lead to political instability.

Contingent extremism is also informed by the literature on 'impersonal influence'—that is, influence derived from individuals' perceptions of anonymous others' attitudes, beliefs, or experiences (Mutz 1992). Noelle-Neumann (1974) theorized that people's propensity to express an opinion about a topic is a function of how they perceive the aggregate distribution of opinions on that topic (i.e., the 'climate of opinion'). According to her 'spiral of silence' theory, individuals subconsciously scan their environments to evaluate the climate of opinion, in order to develop a quasi-statistical sense of where they stand on sensitive issues compared to others. This quasi-statistical sense serves as a social adaptation; when individuals express an opinion they perceive to be unpopular, they risk social isolation and loss of status, and might restrain from expressing it.

A person's fear of isolation is triggered by the belief that others will consider him not merely mistaken but immoral for his opinion (Noelle-Neumann 1989). In other words, the spiral of silence is only activated for morally laden or socially sensitive issues, because people feel no intrinsic need to monitor opinion climates on issues that carry no social or reputational risk.

Using experiments embedded in national telephone surveys, Mutz (1989) and Mutz (1992) empirically demonstrated this differential willingness to express opinions under varied opinion climates. However, Mutz asserts that collective opinion influences individual opinions not through normative pressure, fear of isolation, or group identification mechanisms, but through the sheer numerical magnitude of 'others' that support certain opinions or attitudes. According to Mutz's cognitive response model, public opinion cues induce attitude change by prompting people to mentally rehearse arguments related to the issue position: "by means of these cognitive responses, people are essentially engaged in a self-persuasion process whereby their own opinions move in the direction of the reasons that would not otherwise have come to mind" (Mutz 1992, 98).

In addition to the spiral of silence and the cognitive response model, scholars have theorized numerous mechanisms that activate impersonal influence. One line of research suggests that impersonal influence works through the intrinsic desirability of being on the winning team (Bartels 1988; Norrander 1991). This type of motivation has been termed a 'bandwagon effect' in the context of opinion poll influence. A related literature characterizes the phenomenon of impersonal influence as a form of 'cue-taking'—wherein individuals' underlying cognitive motivation is to make the most accurate choice, and public opinion simply serves as a strong cue, or heuristic, indicating the most intelligent choice to make (Chaiken 1987; Bartels 1988).

The spiral of silence model is the most probable mechanism at work in contingent extremism, given a number of scope conditions that make 'cue-taking' and the 'bandwagon effect' untenable explanations. As Figure 2.1 illustrates, this is because far right parties increase their political momentum in a non-linear fashion after surpassing certain minimal thresholds of electoral support. In line with this trend, I expect that contingent extremists begin to support far right parties once they are no longer perceived as 'fringe' minorities. In other words, contingent extremists are comfortable belonging to a political minority—just not a marginal one. This eliminates the bandwagon effect as a possible mechanism by which contingent extremists come to express their preferences openly, because it runs counter to the notion that individuals intrinsically wish to be associated with the winning opinion. Similarly, supporting the far right when the party is still electorally weak would not represent cue-taking toward the modal, or 'correct' public opinion choice.

Mutz's cognitive response model is an alternative to the spiral of silence and lies within the scope of this study. Although the polling experiments offer no direct test of Mutz's cognitive response model, the premise of 'impersonal influence' is that the sheer numerical magnitude of 'others' induces a self-persuasion process by which people mentally rehearse arguments related to the issue position. Two central premises work against the this framework in understanding contingent extremism: (1) I do not expect expressed party support to increase linearly with higher magnitudes of 'other' far right supporters; and (2) I expect expressed party support to increase when the numerical magnitude of 'others' is still quite low. It is unlikely that receiving information about moderately low party support in polls would activate a cognitive response model, whereby respondents mentally rehearse the possible reasons people have for supporting the far right.

Although my theory of contingent extremism differs from Noelle-Neumann's spiral of silence in some fundamental ways, it provides two analytic advantages as a conceptual starting point. First, far right extremism is one of the most morally laden political ideologies in contemporary politics. More than any other political ideology, the far right's views on issues like race and immigration are perceived and discussed by many others as inhumane. Since the spiral of silence theory was first proposed in 1974, it has replicated most consistently when examining morally laden issues, given that the threats of social isolation and loss of status are most palpable for these issue types (Noelle-Neumann 1989; Scheufele and Moy 2000; Scheufele, Shanahan, and Lee 2001). For example, although the Green party family in Europe is also relatively new and small, I would not expect supporters of green, eco-conscious politics to experience the same social pressure to default to more mainstream contenders, because their viewpoints are not considered immoral in contemporary politics.

Second, the spiral of silence is not derived from the premise that individuals desire to be associated with the statistical majority, or that individual opinions are easily malleable under various types of 'impersonal influence.' Cue-taking, bandwagoning, and cognitive response models all propose that shifts in the direction of the majority opinion are most likely to occur among citizens with low levels of information or little motivation to think about an issue (Petty and Cacioppo 1979; Petty and Cacioppo 1981; Patterson 1980; Kaplowitz et al. 1983; Geer 1988; Geer 1989; Mutz 1992). However, I expect that contingent extremists are politically knowledgeable and highly motivated to think about political issues—their contingency is, in part, derived from their motivation to bring their political preferences in line with mainstream politics.

One important deviation I make from the work of Noelle-Neumann, Mutz, and Kuran is that these authors theorize the public expression of opinions. That is, they theorize when people are willing to express controversial opinions in contexts where this opinion is visible to others. Instead, I focus on the private act of voting, where there are no such reputational costs. I suggest that some extremists have an internal tendency to conform with mainstream values, and this tendency can cause them to either support a more established right-wing party or not vote at all. As information environments evolve to legitimize the far right, this mechanism still produces threshold effects whereby contingent extremists are activated as far right supporters.

Another important contribution I make is to explain why some minority opinion holders are vulnerable to mainstream pressures whereas others are not. I posit that the geographic placement of people with extreme views shapes the degree to which they perceive the far right as non-conformist and fringe. For those in low party support regions, the social stigma of the far right is much stronger than for those from high party support regions.

In sum, people who have far right attitudes and preferences will not vote for a far right party when their information environments discourage viewing these parties as legitimate. The opinion climate acts as a filter that determines which parties are considered serious political options. My theory of party legitimacy thereby dovetails into sociological work on "discursive opportunity structures" which theorize that far right parties gain legitimacy through increased public exposure because it provides opportunities for members of the media to express their support visibly (e.g., Giugni et al. 2005; Koopmans and Muis 2009).

In the following section, I examine drivers of opinion climates, and describe how changes in these drivers help explain the conditions under which contingent extremists become party supporters.

The Drivers of Opinion Climates

Knowledge of the attitudes, beliefs, or experiences of anonymous others must, by definition, reach people through mediated information. Noelle-Neumann distinguishes two primary information sources: (1) immediate social environments (comprised of personal networks) and (2) mass media. The latter can become particularly important in situations where people lack direct contact with a given issue, or more generally, when they have trouble accessing information about the climate of opinion (Scheufele and Moy 2000).

Social environments inform opinion climates in two ways. First, they facilitate direct, interpersonal discussions of relevant issues. Second, they facilitate passive, day-to-day observation of issue opinions, and help people conduct consistent 'quasi-statistical' opinion monitoring (Noelle-Neumann 1995b; Noelle-Neumann 1995a). For far right parties, this implies that breakthrough elections—which usually happen locally and regionally first—generate interpersonal social discussions, especially in regions where party vote share increased. These discussions allow contingent extremists to 'update' their perception of far right party popularity and may activate them as supporters. However, when interpersonal discussion of the far right reinforces their stigma as political pariahs, contingent extremists observe this as well.

The second major driver of the opinion climate is the media, because it serves as the conduit of information about dominant mass opinion. However, the contemporary fragmented media environment also serves as a conduit of information on marginalized opinions. With enlarged choice over where and how to gather news, people engage in selective exposure and follow outlets that reinforce their pre-existing political views while avoiding contradictory information (Mutz and Martin 2001; Stroud 2008; Arceneaux, Johnson, and Murphy 2012).

This filtered media environment facilitates the destigmatization of the far right in a few different ways. First, contingent extremists may opt into right-wing media networks that are less likely to stigmatize far right parties and ideas. Second, biased political coverage makes it more difficult for contingent extremists to compare far right views to the majority, leading them to misperceive extreme parties as compatible with the climate of opinion (Perry and Gonzenbach 1997; Perry and Gonzenbach 2000). Third, right-wing networks differentially promote pre-election polls and other political information that indicate far right popularity.

Opinion Climate Thresholds

When do we expect contingent extremists to support the far right? The social movements literature helps answer this question. Kuran theorizes that each member of a society has a different psychological propensity to protest against the regime. Some are 'relatively insusceptible to social pressure' and are first to protest when they are discontent with their governments (Kuran 1989, 19). The equivalent of the insusceptible protester is the non-contingent extremist, who is willing to support a fringe party in spite of mainstream pressures. For example, we may surmise that in 1997, when only 5% of European voters supported far right parties in parliamentary elections, all party voters were non-contingent extremists, because they made the choice to vote for the far right when it was extremely unpopular to do so.

By definition, contingent extremists have a lower propensity to withstand social pressure. Their conformism correlates with a high destigmatization threshold at which they consider far right parties legitimate political options. Because of the threshold gap between non-contingent and contingent extremists, a ‘fringe far right’ equilibrium with only non-contingent support can self-sustain for multiple electoral cycles.

However, even seemingly stable political equilibria may be vulnerable to a minor change in the distribution of individuals’ internal ‘defection’ thresholds (Kuran 1989). For example, even if just a small portion of contingent extremists are activated by their neighbors, or by their media networks, it can set in motion a series of defections which alter the national political equilibrium. This political shift is apparent in Europe, where the far right’s overall vote share has increased by more than 10% in two decades (Tartar 2017).

Deviating from the literature, I posit that far right supporters’ party ‘defection’ thresholds are unevenly distributed across the population. Whereas game theoretical models of protest often feature an assumption about normally distributed propensities to protest in the population (Yin 1998), certain psychological anchors may exist at which contingent extremist perceive the far right as legitimate. The presence of such a threshold (or thresholds) would help explain the non-linear electoral momentum of Europe’s far right illustrated in Figure 2.1.

In the following section, I derive three hypotheses on opinion climates and how they shift contingent extremists’ willingness to support the party. I test these hypotheses in France, Germany, and Hungary (n=4,776) through polling experiments and observational evidence.

2.4 Theory and Hypotheses on Contingent Extremism

My first empirical task is to establish that contingent extremists exist. I define contingent extremists as proponents of extreme right-wing ideas that are sensitive to the ‘climate of opinion’ and support the far right contingent on the party gaining political legitimacy. Because contingent extremists have an internal tendency to comply with mainstream values and norms, I hypothesize that they are particularly sensitive to one dimension of political legitimacy: the level of public party support. In the polling experiments, contingent extremism is operationalized as the difference between rates of party identification in randomly assigned party polling treatments. This leads to the following hypothesis:

> H_1: *Contingent extremists will not support the far right when they are led to believe that very few others support the party.*

By this logic, contingent extremists will not support the far right in social environments comprised of few far right voters, because these environments signal that the party is illegitimate. I test the effect of local social environments on contingent extremists by matching survey respondents to their electoral districts through their geolocation data. Once I identify each respondent's local social environment, I test whether contingent far right extremism is more prevalent in districts with low or high party vote share. This leads to the following hypothesis:

> *H_2: Contingent extremists will not support the far right when they reside in low party support electoral districts. They will support the far right in high party support districts.*

Experimentally, it is difficult to identify non-contingent extremists' internal thresholds for party support because, by definition, they have already met it. However, for contingent extremists, it may be possible to identify such thresholds. These inflection points may differ slightly for each country due to national norms as well as electoral systems that make the success of the far right more or less likely. This leads to the following hypothesis:

> *H_3: Contingent extremists' party support increases at a threshold (e.g., inflection point) of far right party popularity, rather than following a linear trajectory.*

2.5 Research Design: Survey Experiments

Case Selection

Although contingent extremism is foremost a function of subnational politics and the local social norms they sustain, certain country characteristics fix baseline levels of extremism in the electorate. In other words, country-level variation on institutional and social features helps understand the factors that exacerbate contingent extremism (Scheufele and Moy 2000). This experiment examines Hungary's Jobbik, France's Rassemblement National (RN)[1], and Germany's Alternative for Germany (AfD) as cases that provide broad country-level variation on the following criteria: (1) far right vote share in recent elections; (2) the tenure of the far right party in parliament; (3) the degree to which the governing coalition differs in its political ideology; and (4) the strength of mainstream liberal norms in the country. Given these national differences, I expect survey respondents from each country to perceive the baseline legitimacy of right-wing extremism differently.

One implication of contingent extremism is that far right supporters incur different reputational costs based on the party's vote share. RN's vote share in the 2017 French parliamentary elections was 13.0%, AfD's vote share in the 2017 German federal elections was 12.6%, and Jobbik's vote share in Hungary's 2018 parliamentary elections was 19.1%. Although Hungarian contingent extremists should be most open to supporting the far right based on

[1] Formerly known as Front National (FN), the party was officially renamed Rassemblement National in June 2018 as part of a 'rebranding effort' to signify its support of the nation (Anderson and Vinocur 2018)

this metric alone, a history of successful elections is likely important as well. For example, RN received 4.7 million votes in the 2014 European Parliament election, finishing first with 24.9% of the vote and 24 of France's 74 seats. Moreover, during the 2017 French presidential election, RN's leader, Marine Le Pen, participated in a run-off election against Emmanuel Macron as one of the top two candidates in the first round of voting. Based on electoral history, AfD is the weakest party of the three. I expect stronger social pressure on Germans to support mainstream options due to AfD's weaker electoral performance.

A long party tenure may also increases far right legitimacy. RN has a significantly longer tenure than either Jobbik or AfD, although all three have experienced break-through elections during their tenure. Founded in 1972, RN failed to gain more than 0.5% of French votes for multiple election cycles. In January 1984, the party made its first appearance in a monthly poll of political popularity, with 9% of respondents holding a 'positive opinion' of RN (Shields 2007). Six months later, RN experienced a sudden political breakthrough during the June 1984 European Parliament elections, winning 11% of the vote. Founded in 2003, Hungary's Jobbik experienced a sudden break-through election as well. Jobbik failed to win any parliamentary seats in Hungary's 2006 election cycle but surged to 16.7% of national vote share in 2010. AfD has the shortest political tenure of the three parties. Founded in 2013, AfD participated in the 2013 German federal elections but missed the 5% threshold to enter the Bundestag. By the 2017 federal elections, however, it garnered 12.6% of the votes.

The political ideology of the governing coalition also influences the far right's baseline legitimacy. Mainstream right parties often try to 'outbid' their successful far right counterparts—moving their policies further right to recapture and maintain their vote share (e.g., Adams and Somer-Topcu 2009; Mudde 2007). Outbidding implies that the far right is normalized in part by the governing coalition's ideological borrowing. While in Hungary, this is certainly true for the increasingly rightward moving Fidesz government, France and Germany's centrist governments have not legitimized RN or AfD with similar outbidding strategies. In turn, I expect Hungary's Jobbik party to have more political legitimacy than either RN or AfD.

Finally, a long history of liberal democracy may increase the reputational cost of supporting the far right. In both Germany and France, the liberal democratic order has been stable for many decades. In Germany, there are additional layers of taboo attached to the far right due to the country's historical memory of World War II and German Nazism. The tradition of a liberal democratic order is much shorter in Hungary, and thus, Hungarians' preference for mainstream liberal political options may be comparatively weak. The bounds of socially acceptable party positions, ideology, and rhetoric are also less established in post-Communist EU member states.

Based on far right supporters' reputational costs along these four dimensions, I expect to find the strongest evidence of contingent extremism in Germany, where the AfD's brief party history and low vote share—as well as the country's strong liberal democratic foundations and historical memory—all exacerbate the stigma of the far right. I expect to find less evidence of contingent extremism in Hungary.

Survey Recruitment and Subject Pool

This survey experiment's subject pool is comprised of 1,015 Hungarian, 1,770 French, and 1,991 German Facebook users, recruited through Facebook ads.[2] Although there is a dearth of research on Facebook as a recruitment mechanism, a few scholars have noted that Facebook survey participants, on average, are more demographically diverse than standard Internet samples and significantly more diverse than typical college samples (Boas, Christenson, and Glick 2018). As a result, survey researchers are increasingly using Facebook to recruit respondents (Boas, Christenson, and Glick 2018).

There is also limited knowledge as to what systematic differences there may be (if any) between a Facebook user who is willing to participate in an online survey and one that would not. This is an important limitation of this survey sample. I took two main measures to limit bias in the subject pool. To avoid recruiting respondents with unrepresentatively strong political interests, the ads frame the survey as a simple public opinion poll by academics. To increase the representativeness of the respondents and prevent any major barriers to entry, all surveys and corresponding Facebook ads were hosted in respondents' native languages—Hungarian, French, and German. Table 2.1 summarizes respondents' demographic characteristics. The most noticeable bias is towards young survey respondents (18-24 years old) in the French and German survey samples. The skew towards younger respondents in Germany and France relative to Hungary cannot be explained by the median age in Hungary (41.7), since it is on par with France (41.2) and lower than in Germany (45.9). It is instead a feature of country-level social media consumption. Eurostat data indicate that 83% of Hungarian Internet users between the ages of 16 and 74 used social networks such as Facebook and Twitter in 2016—the highest rate in the European Union (Eurostat 2017). In comparison, just 47% of French and 56% of Germans reported using social networks in the same time frame—and in both countries, the user base skewed very young (Department 2019; Sonnichsen 2019). While it is important to note this bias in age distribution, I expect that age is not highly correlated with contingent extremism.

There is also a skew toward women in the French survey sample. There are two reasons this is not a cause for concern. First, Although women typically support far right parties at significantly lower rates than men, this is not the case in France. Post-electoral surveys indicate that, due in part to Marine Le Pen's ascendance as Rassamblement National's leader, RN has closed the gender gap (Mayer 2015). Second, I do not expect that women are more likely to be contingent in their extremism. I presume that social stigma functions similarly across genders. However, analysis of heterogeneous treatment effects is necessary to confirm expectations around age and gender.

The German survey sample is highly educated—but in comparison to country-level educational attainment statistics, so are the Hungarian and French survey samples. This is likely due to the survey recruitment strategy; college-educated Facebook users may be more interested in participating in a public opinion poll advertised by academics.

[2] The Hungarian survey was fielded in April 2016; the French survey was fielded March-April 2019; and the German survey was fielded April - June 2019.

All other demographic characteristics roughly correspond with national census data. Most survey respondents are either Catholic or non-religious (or Protestant, in the German sample). As expected, the overall survey pool is skewed right-leaning in Hungary, where the right-wing Fidesz party has been in power since 2010, whereas the respondent pool is skewed left in both Germany and France. Urban residents are somewhat over represented in France and Hungary. Despite these imbalances, Table 2.1 represents a highly diverse, and moderately nationally representative respondent pool, adding external validity to the experiment.[3]

Table 2.1: Demographic Characteristics by Survey Sample (in %)

	Hungary	France	Germany
Male	46.9	36.8	48.1
18-24 years	8.3	44.9	49.9
25-34 years	9.6	7.0	13.7
35-44 years	17.0	9.3	7.4
45-54 years	20.2	14.9	11.4
55-64 years	31.0	15.3	10.5
65 years +	13.9	8.6	7.1
less than HS	23.5	11.5	16.2
HS diploma	38.8	35.9	7.3
some college	10.8	33.0	31.1
college degree	14.8	6.9	28.3
graduate degree	12.0	12.8	17.0
Catholic	46.8	24.4	15.8
Protestant	4.4	2.1	17.3
Jewish	0.3	0.7	1.2
Muslim	0.05	2.2	1.3
non-religious	36.0	21.4	56.0
other	12.0	49.2	8.4
politically left-leaning	28.8	52.1	44.0
politically moderate	34.7	27.6	46.3
politically right-leaning	36.6	20.3	9.7
Resident of capital region	34.8	21.6	6.9
Resident of 5 largest cities	41.6	35.7	22.8

[3]Note that due to survey fatigue, 20-30% of survey respondents did not complete each demographic question. I impute all missing control values for supplementary regression analyses using the MICE package in R, which employs predictive mean matching (pmm) to impute all numeric variables (see Buuren and Groothuis-Oudshoorn 2010) and Bayesian polytomous regression modeling to impute categorical variables (see Brand 1999, Chapter 4).

Treatment Design

To ensure a high retention rate for the main experiment, respondents were assigned to the randomized polling question at the survey's outset. The main objective of the polling experiment was to induce different 'opinion climates' with randomly assigned snippets of information about the far right.[4] The experiment treated respondents with varied polling information about far right party popularity. Although this study does not examine the effects of the fragmented media on contingent extremism directly, these experimentally varied snippets of polling information serve as a rough proxy for the types of biased information contingent extremists glean through fragmented media networks.

Chronologically, the design of the treatment is as follows. After the collection of informed consent, subjects were randomly assigned to the experimental question, which varied how the far right party was polling: "Recent national polls indicate that X% of [nationality] voters support [far right party name]." Then, respondents were prompted to answer, "Do you personally consider yourself among this segment of the voting population?" Respondents in Hungary and France were randomly assigned to one of four possible experimental polling treatments, and respondents in Germany were randomly assigned to one of six possible experimental polling treatments. Adding a larger range of polling treatments to the German sample was necessary to examine possible inflection points in far right support with greater precision.

According to hypothesis 1, respondents assigned to polls indicating very limited far right support in the country would rarely self-identify as party supporters. In contrast, respondents assigned to polls indicating broader far right popularity would identify as supporters at significantly higher rates. Lacking a strong theoretical prior for the level of far right polling that might indicate party legitimacy to contingent extremists, I employed a broad range of polling values, and later aggregated these into 'high' and 'low' support conditions based on the experimental results. I varied the public opinion polling of the far right between 5%, 15%, 35%, and 55% in France and Hungary, and between 5%, 10%, 15%, 20%, 30%, and 40% in Germany. This range of polling values correspond to the range of district-level far right party vote shares observed in the most recent parliamentary elections for each country.

In other words, although these parties only capture 13-19% of the national vote share, they are very popular in certain regions of each country, which is why providing a wide polling range was necessary to examine the geography of contingent extremism. Moreover, although the far right does not capture the majority of votes in any Hungarian, German, or French electoral district, adding a 55% polling treatment helps rule out the possibility that the 'bandwagon effect,' 'cue-taking,' or 'cognitive response' theories explain changes in the

[4] I also conducted a list experiment to gauge implicit support for the far right, but the experiment was unsuccessful due to the chronology of the survey design. For more information about the list experiment results and why list experiments are often unsuccessful at gauging implicit support for sensitive topics (and why it was unsuccessful in this particular experiment), please see the overview provided in Appendix Table A.2. Imputing missing values, the list experiment is more consistent with the findings in the polling experiment.

rate of party support.

2.6 Findings: Contingency in Low Party Support Districts

I first employ an observational approach to detect any thresholds at which contingent extremists begin to express far right support, and then proceed to the experimental analysis. Both methods yield consistent results, and indicate that extremists' social environments affect their willingness to identify as far right supporters. For both analyses, I employ district-level far right vote share as my unit of analysis to proxy respondents' local opinion climates, because it is the most granular voting data available in all three countries.

I determine each survey respondent's local opinion climate by matching their GeoIP data (i.e., latitude and longitude) to their district's far right electoral vote share. The survey experiment's host platform, Qualtrics, records GeoIP data for all complete survey responses. However, numerous survey respondents stopped short of completing the last few demographic questions, leaving their GeoIP data incomplete. Fortunately, for each partial respondent, Qualtrics returns an individually identifiable IP address. I prevent the loss of important respondent-level geographic data using a GDPR-compliant IP anonymization approach employed by Google Analytics to approximate users' geolocations without querying individually identifiable IP information. I anonymize respondent-level IP data by setting the last octet of each respondent's IP address to zeros in memory. This last octet of data is what makes a user individually identifiable, to the granularity of street address. By 'zeroing' this octet, I am simply querying city-level data, and complete each respondent's GeoIP data (n=4,776 total) with city-level latitude and longitude information, using the IPinfo API. This GeoIP data allows me to carry out robust tests on the effects of local opinion climates.

Figure 2.2 below maps Hungary, France, and Germany, by district-level party support for the far right in the most recent parliamentary elections. 168 electoral districts comprise Hungary, 96 electoral districts comprise France, and 403 electoral districts comprise Germany. The black dots represent individual respondents' geographic location. To conduct the opinion climate analyses, I leverage district-level far right election results from the most recent national election: the April 2018 Hungarian Parliamentary Election, the June 2017 French Legislative Election, and the September 2017 German Federal Election. Notably, although each country has more recent electoral data from the May 2019 European Parliament elections, national-level election data is more reliable because EU elections have significantly lower turnout than do national elections, which often unduly inflates extremists' vote share.

Observational Analysis of Contingent Extremism

Using the district-level far right election results illustrated in Figure 2.2, I test whether contingent extremists have a party popularity threshold at which they are willing to express

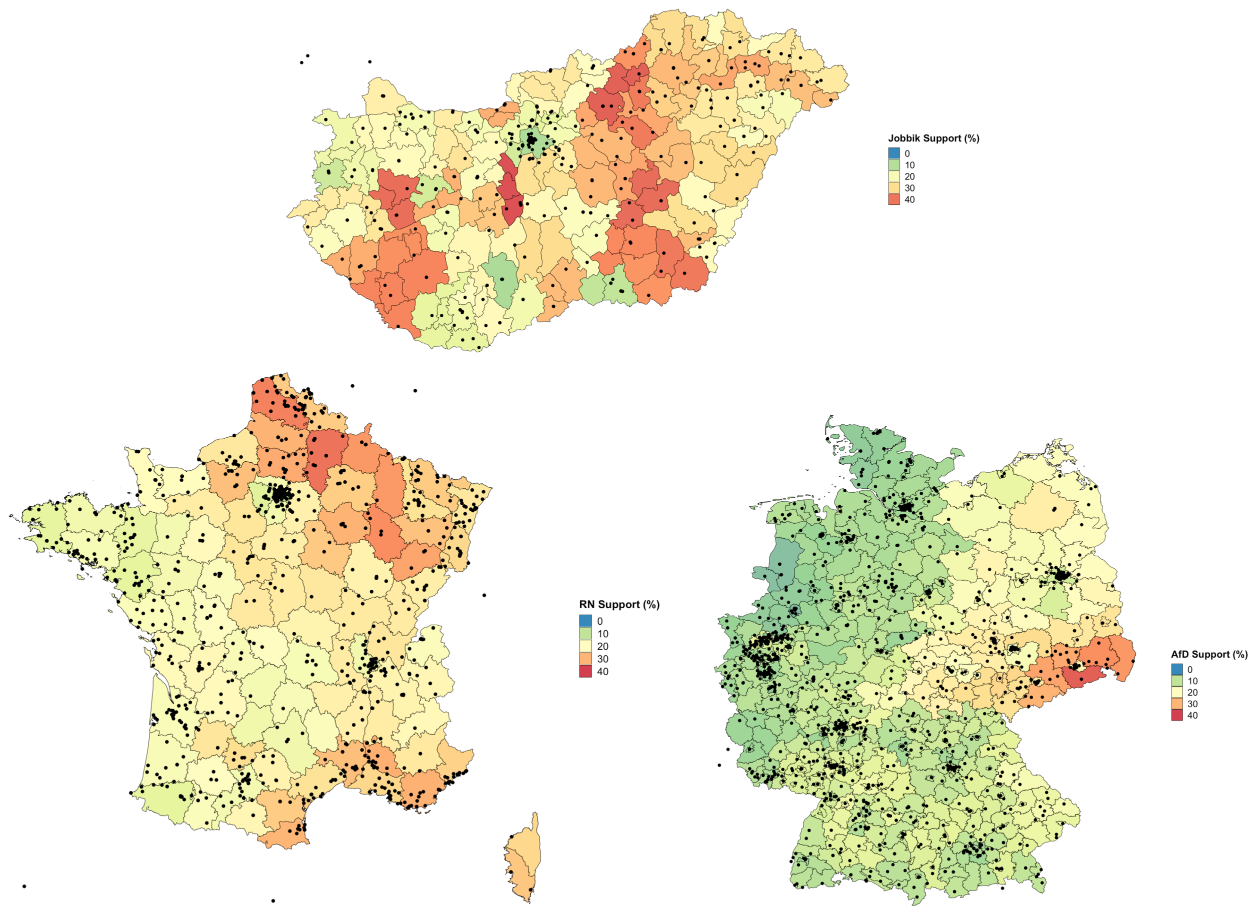


Figure 2.2: Maps of Survey Respondents (n=4,776) and their Local 'Climate of Opinion'

These figures overlay the geographic location of survey respondents with district-level election results from the following elections: for Hungary, the April 2018 Parliamentary Election; for France, the June 2017 Legislative Election; and for Germany, the September 2017 German Federal Election.

support for the far right. To run this test, I first match each survey respondent to their electoral district's far right electoral performance using their GeoIP data. Next, I use responses to the polling experiment to determine whether each respondent supports the far right (1) or not (0). I aggregate these binary responses to the district-level, calculating a far right support ratio for each electoral district. Finally, I map district-level far right election results onto the x-axis as the independent variable, and I map the district-level far right support expressed in the surveys on the y-axis as the dependent variable.

In expectation, the plot should follow an S-shaped (sigmoid) growth curve. If contingent extremists possess thresholds at which they are willing to openly support the far right, there should be an inflection point at which rates of district-level far right identification in the survey begin to increase more rapidly. On either end of this 'inflection zone,' we would also expect to see a smaller slope. This is because, prior to the threshold, we expect that only non-contingent extremists openly identify as far right supporters, and that they do so independent of the local opinion climate. After the threshold is reached, we also expect some gradual leveling off in new support, because even the most socially sensitive contingent extremists defect and support the party in highly favorable opinion climates.

In theory, another potential reality exists. In a hypothetical world where no defection thresholds exist and far right extremists feel comfortable supporting the far right independent of the opinion climate, the data would follow a linear trend (e.g., 45 degree sloped line). We would expect that as district-level far right electoral vote share increased, so would district-level far right support in the country surveys. Figure 2.3 illustrates my theoretical expectation, as well as the alternative possibility.

Figure 2.4 maps district-level far right election results and their relationship to district-level far right support expressed in the Hungarian, German, and French surveys. I employ a LOESS (locally estimated scatterplot smoothing) regression, which is a nonparametric technique that uses local weighted regression to fit a smooth curve through points in a scatterplot. For this analysis, the main advantage of using a LOESS regression is that it can help reveal data trends that are difficult to model with a parametric curve—for example, an inflection point or multiple inflection points in the data. Since LOESS is a very flexible regression approach, it is ideal for modeling complex processes, such as contingent extremism up to a certain party popularity threshold. Figure 2.4's S-shaped curve is consistent with hypothesis 3. Extremists' expressed party support increases at a district-level party support threshold, rather than following a linear trajectory. An inflection point is clear just around one-fifth of district-level far right vote share, and admitted far right support levels off by approximately one-fourth of district-level far right vote share. The most important implication of this result is that far right parties may experience substantial growth in electoral support once the public perceives their popularity as having reached one-fifth of the total voting population.

In the experimental section below, I analyze the polling experiment using the information gained from Figure 2.4. Since certain electoral districts represent 'favorable' opinion climates, whereas others represent 'unfavorable' climates, I expect to identify significantly more contingent extremists in unfavorable electoral districts.

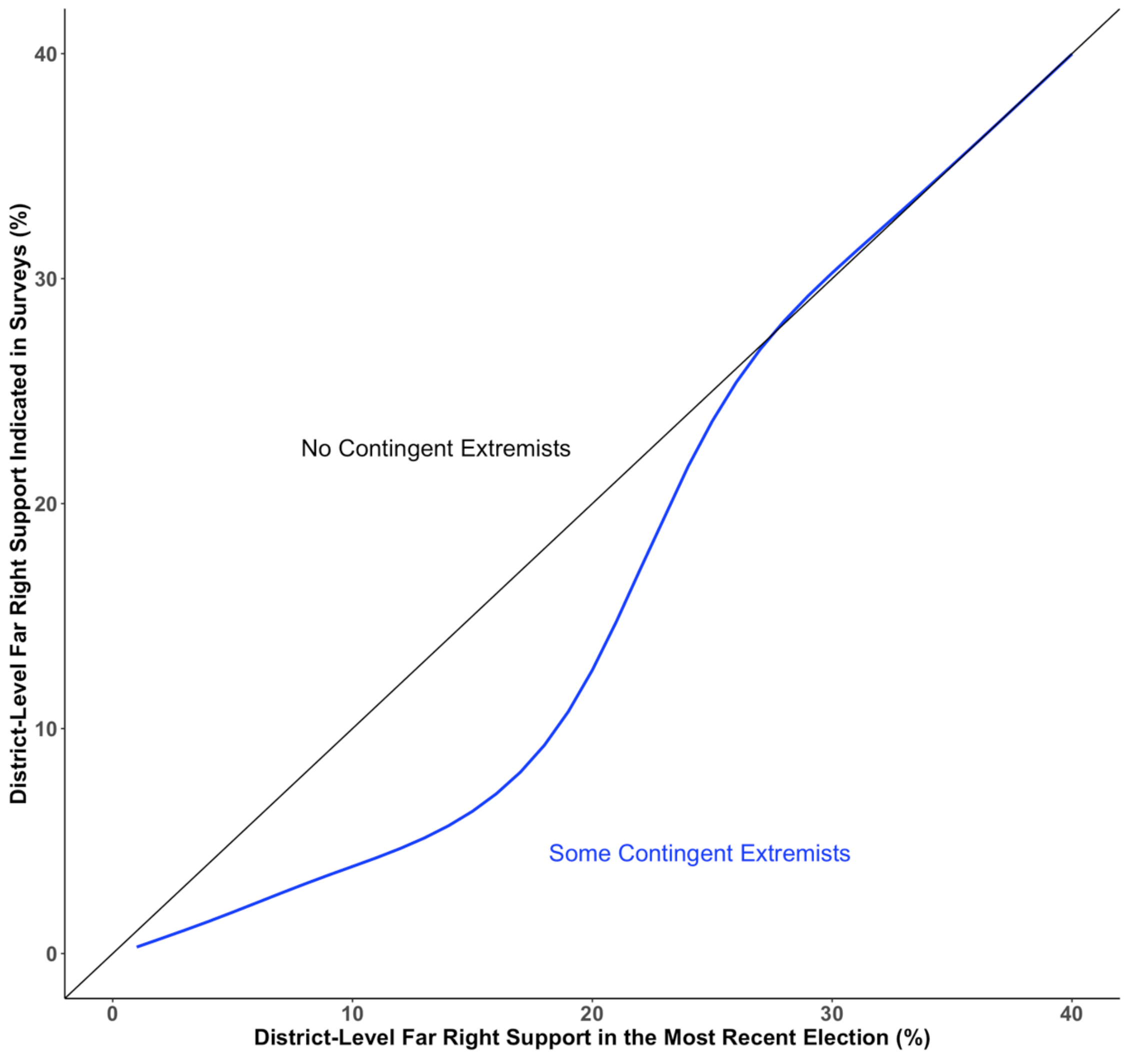


Figure 2.3: Alternative Models of Far Right Support

In a hypothetical world where no defection thresholds exist, we expect that as district-level far right electoral vote share increased, so would district-level far right support in the country surveys (45 degree sloped line). In a hypothetical world where there is a mix of contingent and non-contingent extremists, we would expect to see an S-shaped curve.

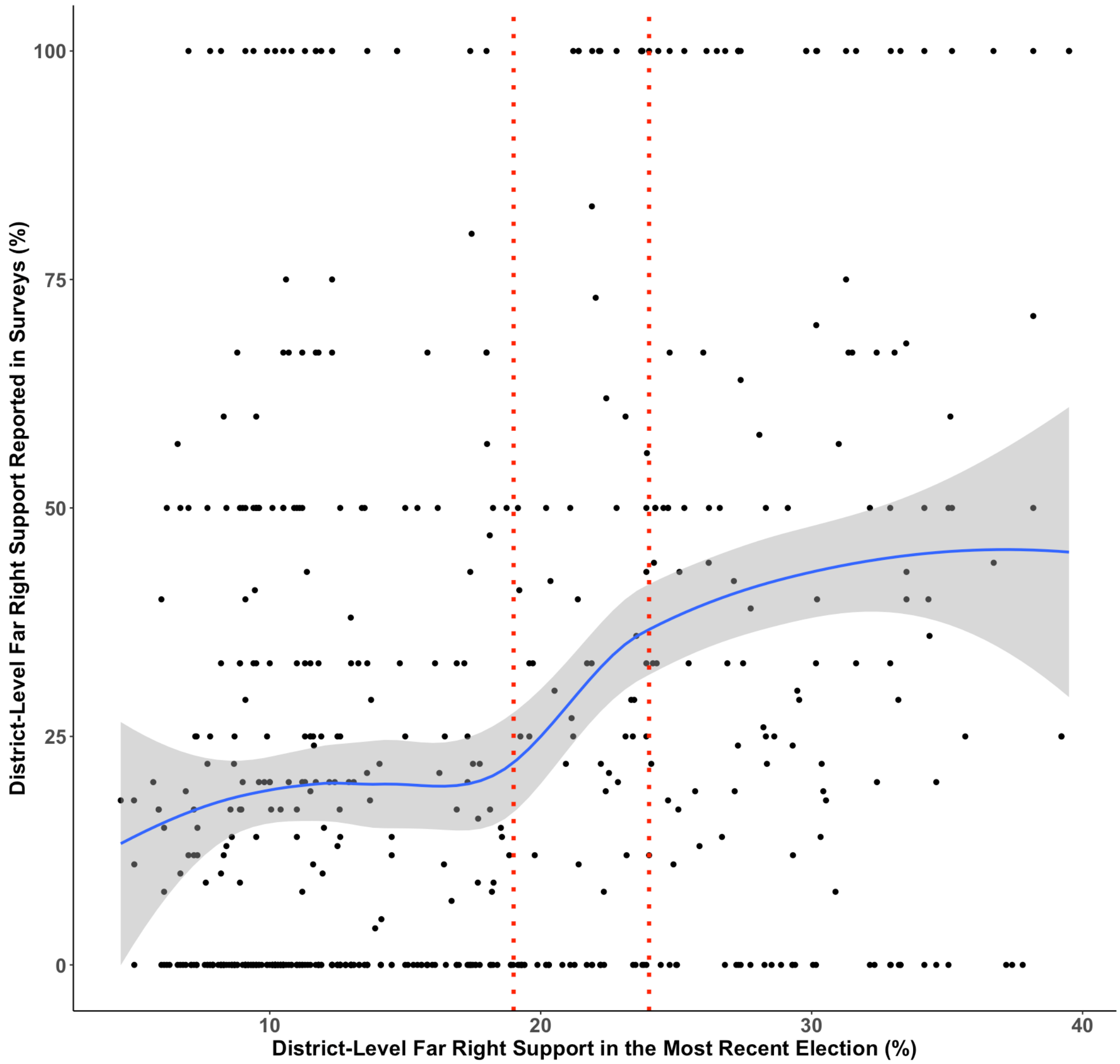


Figure 2.4: District-Level Far Right Vote Share and Admitted Far Right Support in Surveys

LOESS regression mapping far right vote share to rates of admitted far right support in the survey, pooling across Hungary, France, and Germany. Inflection points at 19% and 24% are marked through red dotted vertical lines.

Experimental Analysis

In all three countries, I find that respondents identify as far right supporters at higher rates in the experimental treatment when they are provided with a snippet of information about the party being popular according to recent national polls. Comparing the mean far right support rate in low versus high polling treatment conditions, Hungarian far right support is 39.9%, French far right support is 16.2%, and German far right support is 16.3% in the low party polling treatments. In comparison, Hungarian far right support is 45.1%, French far right support is 20.5%, and German far right support is 23.2% in the higher party polling treatments (see Appendix Table A.1 for a table of these results).

This finding is consistent with Hypothesis 1, and with the concept of contingent extremism more generally. However, the results are more robust in certain cases than in others. When comparing polling treatments in 'high' (30-55%) and 'low' (5-20%) treatment conditions, the differences in rate of far right support are highly statistically significant in Germany (p=0.00014), second strongest in France (p=0.019), and weakest in Hungary (p=0.0954). This finding is in line with the general expectations outlined in the case selection discussion.

Figure 2.5 below illustrates the results of a logistic regression that combines data from all three countries to predict the probability of far right support in each disaggregated treatment condition, grouped by country. In line with the observational finding that contingent extremists have a party defection threshold, there is a clear distinction between the predicted probability of expressing far right support in the 20% polling treatments and below, and in the 30% polling treatments and above.

In addition to this country-pooled logistic regression, I run separate logistic regression models for each country. These country-level polling treatment results are available in Figure A.1 of the appendix. The full logit models for each country can be found in Appendix Table A.3, Models 1, 3, and 5. The data indicate that, relative to the low polling treatments, the high polling treatments increased the probability of far right identification by approximately 5.5% on average. The odds of far right identification are not sensitive to the inclusion of the demographic control variables featured in Table 1.1, as the robustness checks in Appendix Table A.3 Models 2, 4, and 6 show.

Finally, I use the results of the polling experiment to test hypothesis 2—whether contingent extremists reside in social environments with few far right voters. In other words, I test whether snippets of information about the 'climate of opinion' matter more for contingent extremists. Figure 2.4 indicated that expressed far right party support increases between 19 percent and 24 percent district-level party vote share. The inflection points along these curves are roughly consistent across all three countries. As such, in the proceeding analysis, I use 22 percent as an approximate cut-off to classify districts as either socially favorable or unfavorable to the far right. Figure 2.6 below employs the 22 percent threshold to separate each country's results for the polling experiment by the favorability of the local (district-level) opinion climate. In all three countries, in unfavorable opinion climates, numerous respondents support the far right only when presented with high party polling data. This sensitivity signals the presence of contingent extremists. In comparison,

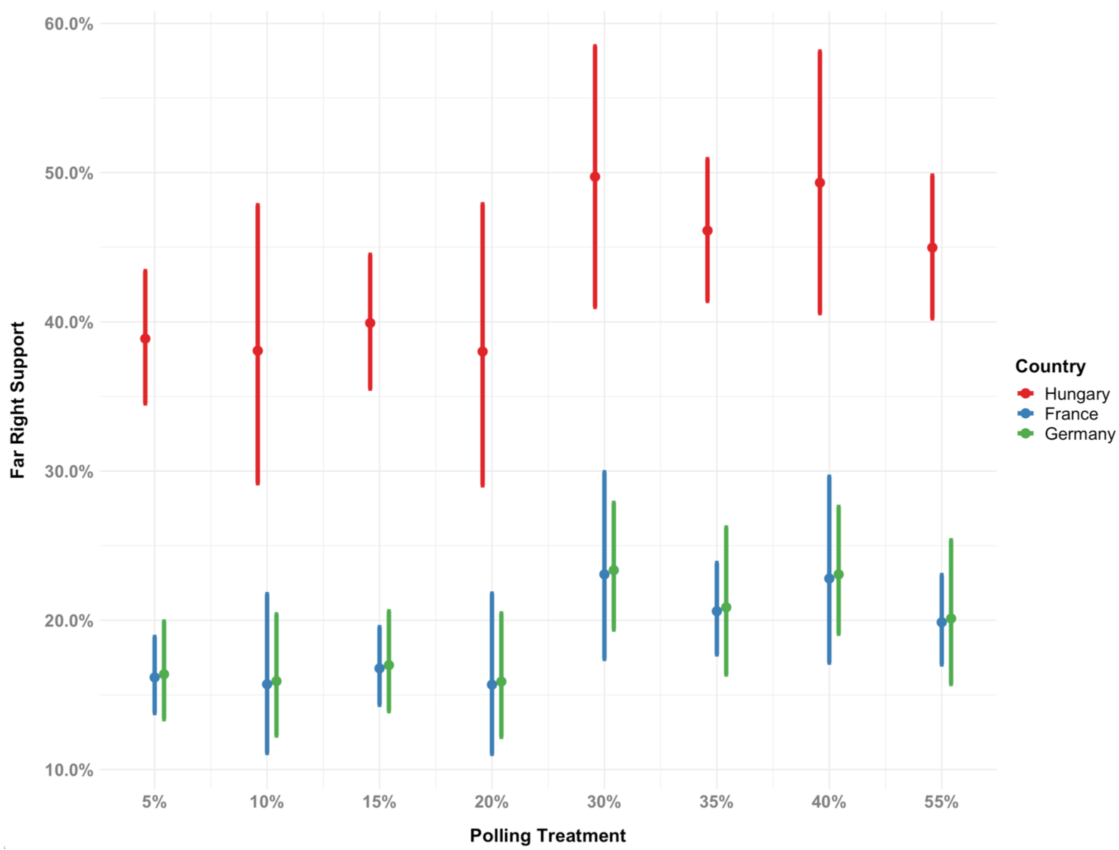


Figure 2.5: Predicted Probability of Far Right Identification by Treatment Condition (with 95% CIs)

Logistic regression predicting far right identification based on treatment condition. Note that the 10%, 20%, 30%, and 40% treatment conditions were unique to Germany, whereas the 35% and 55% treatment conditions were unique to France and Hungary. Only the 5% and 15% treatment conditions are tested across all three countries.

the polling treatments have no discernable effect in the favorable opinion climate districts, because respondents from these regions have already 'defected' to open far right support.

The results in Figure 2.6 are driven by contingent extremists residing in unfavorable opinion climates. The associated p-values of the polling treatment effect decrease from 0.095, 0.019, and 0.0001 in the full set of survey respondents to 0.038, 0.013, and 0.00006 in Hungary, France, and Germany[5], respectively, when accounting for variation in local opinion climate. The full logit models can be found in Appendix Table A.3, Models 1, 3, and 5. These findings are not sensitive to the inclusion of controls, as the robustness checks in Appendix Table A.3 Models 2, 4, and 6 show.

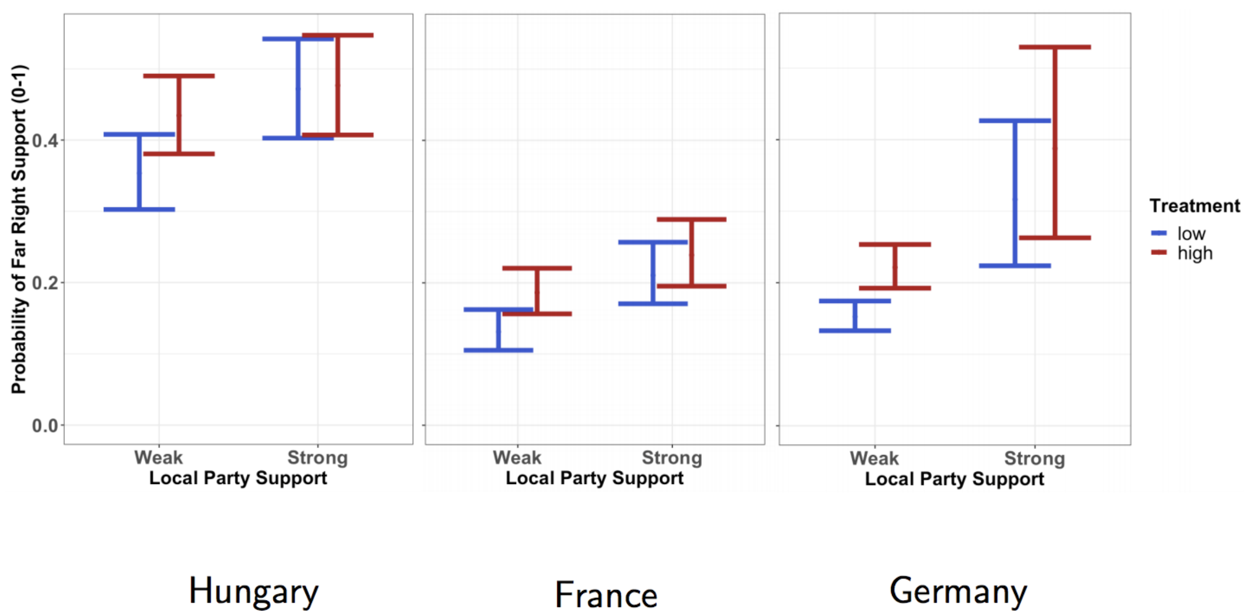


Figure 2.6: Predicted Probability of Far Right Identification by Polling Treatment and District-Level Opinion Climate (with 95% CIs)

Logistic regression predicting far right identification based on district-level electoral support, split by low polling treatments (blue) and high polling treatments (red).

[5]Note that in the German case, fewer than 100 observations are from districts above a 22% electoral support threshold.

2.7 Discussion and Conclusion: Perceived Popularity as a Mechanism for Far Right Momentum

Study Summary

Global trends indicate that far right parties and politicians have reached a critical mass of support in recent years. The main argument advanced in this paper is that reaching this point requires the mobilization of contingent extremists—a segment of supporters that are behaviorally distinct from staunch party loyalists.

Based on a rich behavioral literature on the effects of impersonal influence, I posit that contingent extremists are highly sensitive to the number of other far right supporters in the population. They do not wish to be associated with fringe parties, and they perceive far right support as a reputational risk that they must monitor and socially calibrate. They have two tools at their disposal to do so: (1) their personal social environments and (2) their impersonal media environments.

Using data on district-level far right electoral support, I test the strength of the first 'monitoring tool'—direct social environments. Positing that contingent extremists perceive right-wing extremism as legitimate when the far right is more popular, I leverage district-level discrepancies in far right vote share in French, Hungarian, and German parliamentary elections. As figure 2.6 illustrates, I find that contingent extremism is prevalent in low party support regions. I find no evidence of contingent extremism in regions where the far right already holds substantial electoral support. Using both observational and experimental methods, I also identify an approximate public support threshold, about one-fifth total vote share, at which contingent extremists express party support.

Using a polling experiment, I proxy for the effect of the types of biased information contingent extremists glean through fragmented media networks; namely, different suggestions about party popularity. The polling experiment tests how varied information about the 'climate of opinion' shapes contingent extremists' rate of far right support.

In summary, I find strong evidence that contingent extremists' social sensitivity to mainstream norms influences their willingness to support the far right contingent on the opinion climate. This finding replicates to varying degrees in Germany (n=1,991), France (n=1,770) and Hungary (n=1,015) across varied political systems. When contingent extremists receive snippets of information that suggest that the far right is popular, they are significantly more likely to identify as supporters.

Implications

While many extremists remain contingent, the social stigma of the far right is eroding. The erosion of this norm is at least in part due to a fragmented media landscape, which has produced coverage that helps legitimate the far right. Contingent extremists decided to follow this coverage because it confirmed their priors.

Norm erosion has a local component as well. My study demonstrates that far right legitimacy can be propelled by personal information networks and local social life. In some geographic regions, this process happens much faster than in others. For example, the AfD captures 40% of the vote share in certain rural, eastern German electoral districts and less than 5% in major metropolitan areas. In fact, across all three countries in this study, far right stigma remains heavily balkanized.

This study is comprised of two empirical analyses that show how perceived popularity of the far right activates voters, especially in places where the actual popularity is lower. The most striking finding is that in the absence of actual popularity, perceived popularity can be powerful as a mobilization tool. The strong effect in Germany suggests that information about popularity is effective at mobilizing contingent extremists even in cases where the electoral system makes the success of the far right less likely. In sum, this research demonstrates that information asymmetries have important consequences for political norms and behaviors. Ultimately, these discrepancies can produce meaningful shifts in political equilibria.

Chapter 3

Examining Far Right Campaign Appeals through Digital Political Advertising

3.1 Abstract

The far right is the fastest growing party family in Europe. In 1997, just 5% of the European electorate voted for a far right party in parliamentary elections. Now, 15% of Europeans do. Although party messaging is important to understanding this increase in support, current data on campaign appeals is limited. Cross-national analyses are mostly based on party manifestos, expert surveys, and news media coverage of parties. These data are ideal for investigating formal party ideology, but not how parties prioritize and frame issues in practice. This article is the first to examine party appeals through digital political advertising. Using computational text analysis methods, I evaluate the content of more than 68,000 political campaign ads across 11 European countries and 79 political parties. These ads were fielded on Facebook in the two months leading up to regional, national, and EU elections in the 2019-2020 period. Using this comparative campaign ads dataset, I run three text mining applications to examine prominent theories on Europe's far right. I demonstrate the potential of digital political ad data to address open questions in comparative politics about party ideology and strategy, and then discuss the data's limitations.

3.2 Introduction: The Challenges of Measuring Far Right Party Positions

There has been substantial growth in support for far right parties in European elections over the last two decades. In 1997, just 5% of the European electorate voted for a far right party in parliamentary elections. Now, 15% of Europeans do. Over the same period, the number of European countries with far right parties seated in parliament increased from four to twenty-three. The far right has grown so powerful that in a handful of countries, it has fractured into multiple extreme right parties—each maintaining their own share of seats in parliament.

Party appeals and issue framing, especially during election campaigns, influence voters' political perceptions and vote choice. Campaign messaging is critical to understanding the surge in far right party support across Europe. However, current data on campaign appeals are limited and indirect. Most cross-national studies of the far right examine party manifestos through the Manifesto Research on Political Representation (MARPOR) data (e.g., Harmel and Svåsand 1997; Cole 2005; Jungar and Jupskås 2014; Akkerman 2015). Others rely on expert evaluations of party positions, including the Chapel Hill Expert Surveys (e.g., Rovny 2013; Szöcsik and Polyakova 2019) and the Expert Judgement Survey for European Political Parties (Immerzeel, Lubbers, and Coffé 2016). A few studies examine news media portrayals of the far right and its messaging (Bos, Van der Brug, and De Vreese 2010; Van Spanje and Vreese 2014; Gattinara and Froio 2019).

Through manifestos, expert surveys, and media analyses, scholars have built up knowledge on some important dimensions. Expert surveys suggest that the far right party family is about as homogeneous ideologically as Europe's center-right conservatives and Christian democrats (Ennser 2012). The far right is also ideologically closest to these party families (Immerzeel, Lubbers, and Coffé 2016). Party manifestos suggest that established center-right parties pivot further right in reaction to new far right parties when the established parties experience poor election results (Harmel and Svåsand 1997). Evidence from manifestos and surveys suggest that far right parties have converged ideologically on socioeconomically centrist and socioculturally authoritarian policy positions (Jungar and Jupskås 2014). Media analysis suggests that far right party coverage is overwhelmingly focused on the topic of immigration (Gattinara and Froio 2019).

Systematized knowledge about formal party positions is critical, but it has limitations. Most importantly, far right parties may be incentivized to present a more mainstream ideological profile in formal documents such as party manifestos than in digital ad campaigns. Discrepancies between formal party documents and more informal campaign material may also bias expert evaluations and media reporting. These data are less ideal for understanding how parties prioritize and frame issues in practice.

This article is the first to examine campaign appeals through digital political advertising. Online ads provide unique leverage in understanding far right ideology and messaging in particular. They serve as a more informal—and likely more honest—documentation of far right party views and positions. Moreover, most of these ads are decentralized, since party candidates and local party organizations design and run their own ad campaigns. This feature helps capture variation in messaging and ideology in ways manifestos and other documents cannot.

The remainder of the article reviews the existing data sources used to examine party positions and issue appeals, then introduces online political advertising as a data source and compares it to current approaches. Next, I examine the content of more than 68,000 political campaign ads across 11 European countries and 79 political parties using computational methods. Using this comparative campaign ads dataset, I run three text mining applications to examine prominent theories on Europe's far right. These applications help demonstrate the potential of digital political ad data to address open questions in comparative politics about party ideology and strategy.

3.3 Current Data for Examining Party Positions and Policy Objectives

Evaluating party positions is difficult, in part because formal party documents do not always offer a comprehensive summary of party agendas. Far right parties may be especially prone to signaling their positions less formally, due to pressure to conform to mainstream ideologies. Current comparative party data are thus insufficient to assess far right party positions. The most prominent data is the party manifesto, which a party prepares prior to an election to outline its policy goals. Expert surveys are also commonly used in the literature; country experts rely on a variety of sources beyond manifestos to evaluate party positions, including, for example, parties' track records in parliament and the content of their public speeches. Media portrayals of the far right are also used as data, although less frequently.

The Manifesto Project Database

The most widely used data for estimating party positions and understanding salient policy objectives are party manifestos. The Manifesto Research on Political Representation (MARPOR) database allows scholars to conduct comparative research by providing estimates of dozens of party positions across 56 countries in all democratic elections since 1945. To generate these party positions, the Manifesto Project uses statements up to one grammatical sentence in length as its unit of analysis. Each statement is coded as relevant to one issue dimension. MARPOR codes the number of statements that concern each issue and then divides by the total number of statements in the manifesto to control for manifesto length. The final score for each party on a given issue dimension is the proportion of total statements which fall into the issue category.

There are a number of reasons why a manifesto might not estimate party positions accurately. First, countries may have widely varied ideological distributions. In certain countries, all parties maintain a large spatial distance from each other. If each party's ideology is clear and distinct within the party system, parties might not need to use their manifestos to articulate them in full. Instead, they have the flexibility to present a slightly different policy image or hone in on just a subset of popular issues (Bartolini and Mair 1990). For this reason, far right parties may present more mainstream or moderate party positions than they would in informal discourse (Dinas and Gemenis 2010). The opposite logic holds in party systems where ideological differences are small and less well-defined. In this case, parties—especially smaller ones—look to differentiate themselves on valence issues.

Second, cross-national comparisons may be imprecise due to other country-level variation. As König, Marbach, and Osnabrügge (2013) argue, welfare states and economically liberal states have very different economic policy profiles but may mention certain economic policies at similar rates—creating false ideological equivalence on these measures. As Gabel and Huber (2000) argue, manifestos may also play different roles across countries and cultures. In some cases, the main audience of the party manifesto is the press, in others it is activists, and in others yet it is local party branches. This variation in audience presumably affects the content of manifestos.

Third, certain party characteristics may make manifestos better or worse at predicting policy positions and ideology. As Gabel and Huber (2000) find, mass parties (i.e., centrist or big-tent parties) use their manifestos to touch on a wide range of issue dimensions. However, smaller parties are likely to create manifestos that focus on their signature issues. This may make manifestos a lower quality document, on average, for predicting the ideological profile of extreme or niche issue parties. It is difficult to create a reliable ideological profile for a party when there are many missing values on policy issues in its manifesto. This same logic creates discrepancies between the quality of governing party and opposition party manifestos. Governing parties must address a range of policy issues, whereas the opposition is more likely to emphasize salient issues on which it has 'winning' positions.

Finally, Kitschelt (2007) finds that a number of newer far right parties are missing from the Manifesto dataset. One recent example is Spain's far right VOX party, founded in 2013 and yet to be included in the MARPOR data seven years later. Due to the time intensity of hand-coding manifestos, time lags in the incorporation of smaller, niche parties may be substantial. This limits the usefulness of MARPOR for timely, systematic studies of the far right.

Expert Surveys

Three large expert surveys are commonly used to understand the European political system, and far right politics in particular. The Chapel Hill Expert Survey (CHES) provides estimates of political party positions for a majority of parties in Europe, with individual survey waves from 1999-2017. The Expert Judgement Survey for European Political Parties (2000, 2010) consists of dozens of party characteristics scored across 38 European countries. Although CHES provides greater temporal variation, the Expert Judgement Survey is particularly useful for scholars of the far right because it contains a number of relevant party policy evaluations such as: "immigration restrictiveness, populism, authoritarianism, nationalism, and (anti-)establishment image" (Immerzeel, Lubbers, and Coffé 2011). Benoit and Laver (2006)'s party policy in modern democracies expert survey, which spans 47 countries and 387 political parties, is also widely used in the literature. Although this expert survey is very detailed and features country-specific as well as general questions, its usefulness is limited as a single wave (2003-2004) survey.

Expert surveys typically provide more comprehensive and precise party profiles because country experts develop their evaluations based on numerous sources, including manifestos, speeches, party voting records, and media reports (Benoit and Laver 2006). Experts are also able to incorporate new issues and weigh issue salience in making their evaluations (Huber and Inglehart 1995). Pre-defined scales limit the ambiguity of interpreting final results, minimizing the risk of ad hoc interpretations. Given these features, expert surveys provide a useful benchmark to evaluate the validity of alternative measures (Benoit and Laver 2006).

However, the comparability of expert data is problematic due to ideological variation between countries. For example, a center-left economic view in a welfare state would differ considerably from a center-left view in an economically liberal one. The ideological distribution of parties may also be considerably wider or more restricted across different party systems, making political extremes hard to compare. These types of variation make the aggregation and cross-national comparison of party positions problematic (Bakker et al. 2014). Experts may also have more stable evaluations for certain countries and regions than for others. For example. Benoit and Laver (2006)'s survey indicate significantly higher variation in expert judgments of party positions in post-communist party systems, where experts may have less information on the issue positions of each party simply because they are newer democracies (Huber and Inglehart 1995).

In addition, experts may be biased in their evaluations of certain party families. For example, Benoit and Laver (2006) reported systematic bias in experts' policy placement of far right parties; they rated these parties as more extreme on various issue dimensions than the parties actually were. Using Benoit and Laver's data, Curini (2010) found that experts' own political preferences also biased their placements of center-right and right-wing parties; almost 16 percent of expert party evaluations were imprecise.

Bias in country evaluations may also stem from experts' tendency to view political systems in a coherent and organized manner. For example, experts may incorrectly assume that party families are homogeneous on a number of core policy dimensions, or that there is coherence in how a party approaches two related policy issues. This assumption is especially likely to overestimate the coherence of the far right and far left, since their positions do not fit neatly on the traditional left-right ideological spectrum. This bias may account for why parties in the same party family are evaluated homogeneously on most policy dimensions in expert surveys (Baumann and Gross 2016).

Political Coverage in the Media as Data

Over the last decade, a growing number of analyses honed in on the media coverage of far right parties. The Comparative Campaign Dynamics (CCD) project offers the most comprehensive dataset on the media coverage of electoral campaigns in Europe to date. There are two election cycles in ten countries represented in the dataset. CCD data are based on the content analysis of election-related coverage in the highest circulating center-left and center-right newspapers in the month prior to each election. The analysis is organized in three categories: (1) parties' own discussion of their positions, (2) parties' commentary on other parties' issue positions, and (3) journalists' framing of party positions (Debus, Somer-Topcu, and Tavits 2016).

There are many advantages to focusing on political media commentary. First, whereas manifestos and expert surveys typically provide a single, static measure of party positions for any given election cycle, media content is continuous and can capture the dynamics of an election campaign. Second, in comparison to manifestos, political parties cannot fully control the media narrative (Helbling and Tresch 2011). This reduces certain forms of bias that arise from party agenda setting in formal documents and may lead to a more honest estimation of party positions. Third, news media analyses are able to extract important emotive dimensions to parties and their campaigns. A handful of studies have used news media data to examine parties' negative campaigning and negative issue frames (e.g., Hansen and Pedersen 2015; Elmelund-Præstekær and Mølgaard-Svensson 2014; Pedersen 2014). Fourth, media data reflect issue salience more accurately than expert surveys (Helbling and Tresch 2011). Relatedly, mass exposure to certain issues over others influence how party issue positions are weighed and perceived on election day.

However, there are a few major disadvantages to using mass media content as data on party positions. First, the use of mainstream, high circulation news outlets as data may create systematically biased accounts of far left and far right party positions. Second, as Hellström and Blomgren (2016) observe, the news media tend to stress positional differences between political parties and are unlikely to report on party similarities. We therefore expect that media content analysis will produce party position estimates that are farther apart than true inter-party differences. Third, the news media only publish on matters that are deemed 'newsworthy' (Ruedin 2013). This creates substantial bias in which parties are covered in the news and the frames applied to them. The issue of newsworthiness also implies that certain political topics receive intense coverage across multiple media outlets, from which it is then easy to derive reliable policy position estimates, while other topics receive almost no coverage, from which it is impossible to derive meaningful position estimates.

3.4 Digital Political Advertising as Text-as-Data

This article is the first to examine party positions and appeals on these positions through digital advertising. I collect data through the Facebook Ad Library API, which opened for public release in March 2019 and documents all political advertisements hosted on the platform, as well as limited metadata for each ad (e.g., the name of the ad buyer, the number of ad impressions, total ad expenditure, and audience gender and age demographics). Initially, the API exclusively featured ads run in the United States but expanded to Europe later in 2019. Google and Twitter soon followed in opening political ad archives; these data can be applied to examine similar questions or cross-validate results.

Like other manifesto and media-based position estimates, the digital ads approach assumes that parties' relative word usage provides information about their placement in a policy space. The main advantage of ads in this regard is that each party fields, on average, hundreds of unique online ads in the months leading up to an election. The sheer volume of political ad text available means that it is more feasible to construct reliable ideological profiles for small parties. Relatedly, we expect that there are fewer 'missing values' left on policy issues when the data is comprised of months of ad content than when the data is a party manifesto. Since online ads are time-stamped and contain geographic targeting information, they can also be used to create time-series issue position estimates both sub- and cross-nationally.

Another feature of online political advertising is that parties do not exist as centralized, sole advertisers on Facebook. Each party's main Facebook page is typically responsible for a substantial share of party advertisements, but on average, more than half of advertisements are fielded by party candidates in their own interest. Moreover, regional party organizations (e.g., "AfD-Fraktion Hamburg") often field their own advertisements and focus on content that is most salient to the region. Decentralized issue appeals capture variation in party messaging and ideology in ways manifestos, expert surveys, and media analysis cannot. Decentralized party messaging may be especially critical to understanding how the far right appeals to voters, since these parties typically create electoral footholds in certain regions before seeing a rise in support in others.

In addition to these structural advantages, digital ads are particularly well suited to study the far right, since extreme parties are incentivized to moderate their image in formal, public communications. The far right's targeted political messages provide the most candid account available of what extreme parties communicate to potential voters when ad content is neither publicly visible nor regulated. These features create a clear distinction between the content of online advertisement data and traditional media data. As Benkler, Faris, and Roberts (2018) note, journalists and advertisers for mass media act "in the public eye"—meaning that they have to anticipate reactions to messages from a broader group than just their target audience (372).

There are also a few important limitations to digital ads as data. First, parties advertise online on a voluntary basis, which means that some opt out. While this is uncommon, it restricts the universe of political campaigns available for analysis. For example, this study excludes France as a case because the far right Rassamblement National did not advertise online in the two months leading up to the May 2019 European Parliament election. Second, the Facebook Ad Library is relatively new, and most of the ads hosted are from late 2018 onward. This means that currently, there is at most one election cycle of ad data available to analyze for each country. Third, numerous ads are simple reminder messages to vote in the relevant election or attend a campaign rally. These messages lack substantive content about party positions and may add noise to computational text analysis.

Acknowledging these limitations, there is considerable potential for online ad libraries to provide novel information about party positions and strategy. Moreover, as the ad library grows in real time—with dozens of new ads added each day—so does its potential to serve as a data source of unparalleled size.

Creating a Comparative Campaign Ads Dataset

For researchers looking to create a custom comparative campaign ads dataset through the Facebook Ad Library API, the first step is to select the range of countries of interest. This study examines 11 European countries which have far right parties seated in parliament and held regional, national, or EU elections in 2019 or in early 2020: Sweden, Denmark, Belgium, Germany, Austria, Hungary, Spain, the Czech Republic, Slovakia, Estonia, and Finland. I initially planned to examine three additional countries: France, Greece, and the Netherlands. However, the far right in each country fielded an insufficient number of online ads to make computational analysis feasible. As previously noted, this is an important limitation of the Ad Library approach to political analysis.

Once these country-specific data are retrieved via the API, the next step is to select the period of interest and subset to the ads that were fielded in that time frame. This study examines the two months leading up to each election—although a post-hoc temporal overview of ad start dates indicates that most parties do not begin advertising until approximately 6 weeks prior to the election.

The following step is to subset to relevant political ad creators, since the data retrieved via the API contain hundreds of ad creators that are not campaigning for elections. In this study, I included up to seven political candidates and party regional organizations that fielded the highest number of ads for each party. I did not use a strict ad volume cut-off for inclusion in the dataset, because country and party size affect what may be deemed a considerable number of ads.

Since the computational text analysis tools used in this study were developed and validated using English language text, ad translation was a necessary step in data preparation. Once pre-election advertisement datasets were compiled for each country, I translated all ads (n=68,000) from their native languages to English. I used the Google Cloud Translation API, which is an expensive API to access (costing $20 per million characters to translate) but provides much higher quality translations than the generic Google Translate function. The Google Cloud Translation API can either automatically detect the source language of the input or allow researchers to manually define it. The automatic source language detection feature is critical for working with ad data from multilingual countries (e.g., Belgium).

The automated cloud translation strategy carries the inherent risk that ad translations are of low quality—especially for rare languages. However, manual translation checks on each country dataset indicate that the API's translations are accurate, with the exception of the mistranslation of certain slang in unique languages, including in Hungarian and Estonian. These mistranslations did not negate the overall meaning of the messages in manually reviewed ads. In sum, the threat to the validity of most text analysis applications is minimal. Once translations are complete, the corpus of ad data can be used for numerous comparative text analysis applications. In the next section, I demonstrate the data's usefulness in three distinct applications. These applications help demonstrate how comparative political ad data can be used to (1) create valid point estimates of party positions along a salient issue dimension; (2) quantify the tone and sentiment of different parties' appeals; and (3) examine issue ownership in political campaigns. I use these applications to evaluate three prominent theories on Europe's far right and ultimately argue that comparative ad data has the potential to address open questions in comparative politics about party ideology and strategy.

3.5 Applications of Comparative Campaign Ads Data

There are a number of theories that explain how far right parties appeal to European voters and the conditions under which they maintain an electoral foothold. However, these theories have not been examined through original campaign data at scale. Using more than 68,000 campaign ads across 11 countries, I evaluate three prominent theories on far right party strategy.

First, many studies focus on the grievances that create the initial demand for the far right, including influxes of immigrants (Swank and Betz 2003; Norris 2005; Ivarsflaten 2008; Rydgren 2008; Hainmueller and Hopkins 2014). I use sentiment analysis to examine whether far right parties continue to use grievance appeals to mobilize support after the initial crisis subsides. Then, using a structural topic modeling approach, I examine whether the far right maintains issue ownership on immigration relative to other parties. Third, using an established unsupervised scaling model of document positions, I examine the theory that far right parties maintain an electoral foothold when other political parties within the policy space fail to occupy positions far enough to the right on salient issue dimensions—in this case, on immigration (Kitschelt and McGann 1995; Carter 2005; Norris 2005; Arzheimer and Carter 2006).

Using this range of applications, I demonstrate how the comparative campaign ads dataset can provide unique analytic leverage in evaluating theories of far right politics.

Application 1. The Politics of Grievance

Numerous studies focus on the grievances that create the initial 'demand' for far right parties—including severe economic recessions (Lubbers, Gijsberts, and Scheepers 2002; Bustikova and Kitschelt 2009; Vlandas and Halikiopoulou 2019) and influxes of migrants (Swank and Betz 2003; Norris 2005; Ivarsflaten 2008; Rydgren 2008; Hainmueller and Hopkins 2014). This literature posits that those who feel fearful, angry, or isolated due to rapid social and economic changes are mobilized by far right parties and their promises to return to the previous system through strong law-and-order institutions (Minkenberg 2000; Mudde 2007).

But what happens when these crises subside? Demand-side factors cannot explain, for example, how Europe's far right expanded its base in the years of economic recovery following the 2007-2008 financial crisis. Nor can they account for the surge in far right support since 2017, when the Syrian refugee crisis largely abated. Supply-side analyses that focus on structural factors and party characteristics are a necessary supplement.

In this application of the comparative campaign ads dataset, I run two simple sentiment analyses to examine whether far right parties conduct what may be described as supply-side 'grievance' campaigns. I define grievance campaigns as negative campaigns that focus on—and often exaggerate—social and economic problems, while framing other parties and institutions as responsible for them. Certainly, all political parties make some use of grievance-based appeals. However, given the 'demand-side' origins of far right political contenders, it is plausible that the far right uses this strategy at disproportionate rates as they look to maintain their electoral foothold after crises subside.

To measure the degree to which a party conducts a grievance-based campaign, I calculate the ratio of negative to positive terms that comprise the corpus of their ad campaigns. To do so, I use the Lexicoder Sentiment Dictionary (LSD), which was developed specifically for the sentiment analysis of political text. LSD is optimized for finding sentiment-indicating words in text documents and has the added feature of being able to capture grammatical constructions such as negated positive sentiment and double negative sentiment (Young and Soroka 2012). Since some languages use double negatives frequently in language construction, while others do not, this is an important feature in the context of translated ad campaigns.

The LSD method improves in accuracy with longer text inputs (Young and Soroka 2012). Therefore, rather than relying on the sentiment classification of individual ads, I aggregate all ads to the party level and apply LSD to the resulting party ad corpora. Then, I simply calculate the ratio of negative to positive terms that comprise the corpus of each party's complete ad campaign, disregarding the large number of terms that are neither counted as positive nor negative by LSD.

The results largely validate the idea that the far right conducts grievance-based campaigns at higher rates than all other parties. In seven out of eleven sampled countries, the far right outranked all other parties in their share of negative ad content, sometimes by a large margin. In the most extreme cases—including Belgium, Germany, and Estonia—the far right outranked the closest parties by 15-20% in their share of negative content. In two additional cases—Austria and Finland—the far right ranked as a close second to a left-wing populist anti-corruption party (Austria's JETZT) and a left-wing eco-socialist party (Finland's Vasemmisto). Figures 3.1-3.3 illustrate the distribution of negative ad content across parties in these nine cases.

The far right is also associated with specific emotive appeals, most prominently with fear and anger (Vasilopoulos et al. 2019). More broadly, social conservatism is associated with disgust sensitivity (Inbar et al. 2012). To examine whether the far right uses these emotive appeals at disproportionately high rates in their advertising, I calculate the proportion of text in each party's ad corpus that is associated with anger, fear, and disgust. To do so, I use the NRC Word-Emotion Association Lexicon, which contains a crowdsourced list of more than 25,000 words and their associations with a number of emotions (Mohammad and Turney 2013). Table 3.1 displays the results for far right parties. Check marks indicate that the percentage of terms associated with a particular emotion is higher than for any other party. I find that in 3 out of 11 countries, the far right party uses terminology associated with anger more frequently than any other party in its advertising. In 7 out of 11 countries,

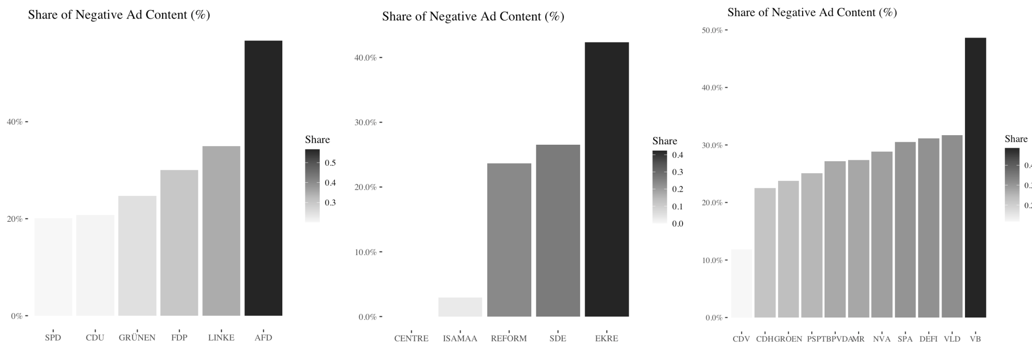


Figure 3.1: Share of Negative Ad Content (%), from left to right: Germany, Estonia, Belgium

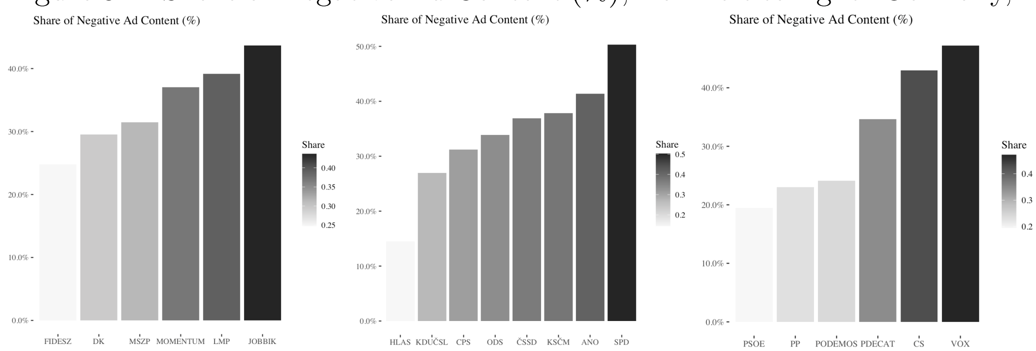


Figure 3.2: Share of Negative Ad Content (%), from left to right: Hungary, the Czech Republic, Spain

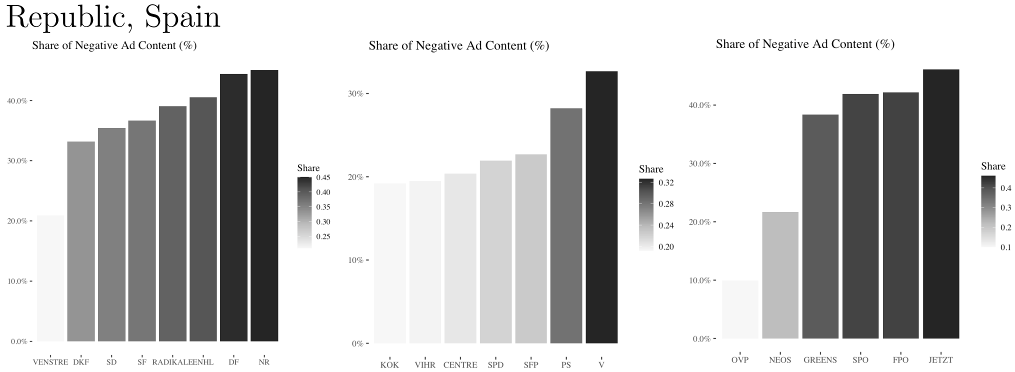


Figure 3.3: Share of Negative Ad Content (%), from left to right: Denmark, Finland, Austria

These figures illustrate the share of negative ad content (%), defined as the proportion of negative terms to positive terms (number of negative terms divided by the total number of 'sentiment' terms), based on the Lexicoder Sentiment Dictionary (LSD). In 7 cases—Germany, Estonia, Belgium, Hungary, the Czech Republic, Spain, and Denmark—the far right outranks all other parties in the share of negative content that comprise their ad corpus. In 2 additional cases—Austria and Finland—the far right ranked as a close second to a left-wing populist anti-corruption party (Austria's JETZT) and a left-wing eco-socialist party (Finland's Vasemmisto).

Country	**Anger**	**Fear**	**Disgust**
Austria (FPO)	4%	5% ✓	2%
Belgium (VB)	7%	5% ✓	2%
Czech Republic (SPD)	5% ✓	6% ✓	3% ✓
Denmark (DF)	6% ✓	6%	2%
Denmark (NR)	6%	7% ✓	3% ✓
Estonia (EKRE)	4%	5% ✓	3% ✓
Finland (PS)	3%	2%	1%
Germany (AFD)	4% ✓	6% ✓	2% ✓
Hungary (JOBBIK)	5%	4%	2% ✓
Slovakia (SNS)	3%	2%	1%
Slovakia (SMERODINA)	3%	2%	1%
Spain (VOX)	6%	6% ✓	3% ✓
Sweden (SD)	4%	3%	1%

Check marks indicate that the % of associated terms is higher than for any other party

Table 3.1: NRC Emotion Lexicon Estimates of the Percentage of Far Right Ad Content that is Associated with Anger, Fear, and Disgust

it uses fear-related terminology more than any other party. In 6 out of 11 countries, it uses disgust-related terminology more than others.

While the far right uses anger-related terminology at moderately high rates, there are other parties—most often left-wing populist—that use anger-related terminology at similar or higher rates. However, far right advertisements clearly feature fear and disgust associated terminology at higher rates than other parties. A number of studies in political psychology find that feeling fearful encourages information seeking, which in turn encourages persuasion (Redlawsk, Civettini, and Emmerson 2010; Albertson and Gadarian 2015; Vasilopoulos 2018). Emotive appeals to fear may be useful as the far right looks to expand their base and persuade moderate right-leaning voters to consider more extreme policy positions. Meanwhile, research suggests that when people are made to feel disgusted, they report more negative associations toward the groups linked to the disgusting behavior (Inbar et al. 2012). The far right may evoke a feeling of disgust toward migrants or other out-groups in their ads, but supplemental analyses would be necessary to confirm this tentative assumption.

Collectively, these results suggest that the far right mobilizes on grievance appeals—and that they do so at much higher rates than other party families. They also use terminology associated with fear and disgust to persuade voters. This application demonstrates that online campaign ads and their emotive features provide leverage in understanding ‘supply-side’ grievance and the mobilization of political discontent. Manifestos, expert surveys, and other formal documents are not suitable for these types of analyses.

Application 2. Far Right Issue Ownership on Immigration

A large literature examines how certain parties come to be associated with particular issues and establish 'issue ownership' over time. One frequently studied relationship is between the rise of the far right and the issue of immigration (e.g., Smith 2010; Dennison and Goodwin 2015; Nygård and Kuisma 2017). As the literature argues, issue ownership develops when an issue entrepreneur politicizes a "previously non-salient event, policy issue or societal conflict, and attempt[s] to gear up public attention over this controversy" (De Vries and Marks 2012, 7). The rise of the far right, then, is attributed to its prioritization of immigration as a primary policy issue.

Issue ownership by political entrepreneurs may be fleeting and temporary, because other parties will aim to integrate salient public issues into their existing ideological profiles to minimize electoral risk (Meguid 2005). This is feasible for them to do because immigration is comprised of a "complex bundle of loosely related policy issues" from which they can selectively integrate ideologically consistent ones (Gattinara and Morales 2017, 279). For example, parties may choose to address border control issues and focus on the perceived connections between security and migration. They may instead focus on integration programs and highlight the potential mismatch between immigrants and national cultural values. Meanwhile, leftist parties often frame the issue as one of human rights and social justice. In theory, as long as parties address immigration consistently, they appease public pressure to 'confront' the issue and erode far right issue ownership.

In practice, however, this is a challenging and resource-intensive task. Mainstream parties may prefer to make small symbolic gestures toward addressing the immigration policy space without leveraging it as a point of mobilization. This may occur for one of two reasons. First, a party may believe that their ideologically consistent immigration 'pitch' is not resonant with a large segment of the population. In this case, the party may prefer to concentrate its limited resources on winning positions. Second, parties may believe that it is bad political strategy to devote attention to immigration, because it legitimizes the far right and increases issue salience.

The choice parties ultimately make in response to the far right is not apparent from currently available data. Party manifestos may address immigration and a host of other topics they consider valence issues out of formality. Media analyses will be equally uninformative on the question of issue ownership. Journalists report on salient, hot-button issues, and prompt parties to make statements on immigration. These statements may simply represent a 'symbolic' party response. Experts rely on both manifestos and media narratives to construct their policy evaluations. Although they also account for the content of public speeches and campaign rallies, expert point estimates are not designed to evaluate party ownership.

Since political parties selectively emphasize the policy issues they "own" (Van der Brug 2004), political advertising is the optimal data to evaluate issue ownership. If parties believe that an emphasis on immigration is advantageous, this will be reflected in their campaign material. If not, they will focus on appealing on their winning positions. In turn, parties' online ads can help discern between two possibilities: (1) the far right devotes a significantly larger share of their campaigns to messaging on immigration than other parties—maintaining issue ownership, or (2) other parties adopt immigration as a central policy priority and emphasize it in their appeals—eroding issue ownership.

In this second data application, I use a structural topic modeling approach to determine whether the far right maintains issue ownership on immigration. I use the Structural Topic Model (stm) R package, which provides tools for machine-assisted processing of text. A uniquely powerful feature of stm is that it allows researchers to estimate a topic model that includes document-level metadata (Roberts, Stewart, and Tingley 2014). In this application, I use metadata on which party fielded each text to examine differences in topical prevalence.

In topic modeling, each document (e.g., party ad corpus) is modeled as a mixture of multiple topics. Topical prevalence captures how much each topic contributes to a document. Because different documents come from different sources, the model allows topic prevalence to vary with metadata about the source. In this application, prevalence is a function of source party, which I code as either 'far right' or 'not far right' by pooling all non-far right party ads in a given country. My objective is to estimate the mean difference in topic proportions for far right parties and all other parties. This helps determine which topics are more prevalent in far right ads and which are more prevalent in others.

There is no specific number of topics appropriate for modeling a given text corpus (Grimmer and Stewart 2013). However, there are a few dimensions on which to evaluate different k topic models. I begin this application by estimating 20, 30, 40, 50, 70, and 100 topic STM models and examining the performance of each candidate model. The stm project provides multiple criteria by which to choose the optimal k topics. I focus on optimizing on three of them (1) semantic coherence, which is maximized when thematically related terms (e.g., 'migration', 'asylum') occur together in a topic; (2) held-out likelihood, which assesses how the predictions of the topic model generalize to unseen data; and (3) residuals, which test for overdispersion and indicate whether more topics are needed to reduce the variance. I find that the 30-topic and 40-topic structural models optimize on these criteria. I then use 40-topic model for the topic prevalence estimation to maximize the number of topics I can explore. Appendix B.1, Figure B.1 illustrates an example plot of stm's various diagnostic values (e.g., held-out likelihood, residuals, semantic coherence) by different k topic models.

For each country's topical prevalence estimation, I plot the change in topic proportions shifting from far right documents to other party documents. Notably, the point estimates associated with far right topics will be larger than the point estimates of other party topics, because there is much greater heterogeneity in the ad content of all other parties combined than of just 1-2 far right parties combined. The objective of this topic modeling application is not to create valid point estimates for different topics, since that would require the comparison of more equally sized documents and more comparable entities (e.g., one party to one party). Rather, my goal is to simply examine the most prevalent issues discussed by the far right in comparison to other parties, to gauge whether there is disproportionate emphasis on immigration in far right campaign ads or immigration topics are prevalent across the political spectrum.

Figure 3.4 provides a cleaned example of these topical prevalence plots. The top 'FREX' words are displayed, meaning words which are both frequent in and exclusive to a topic.[1] I do not display all 40 topics since many of them are get-out-the-vote themes or simple reminders about upcoming campaign rallies, televised debates, or candidate townhalls. I display the topics that are clearly centered on political issue themes. As the Hungarian sample illustrates, it is not always straightforward to interpret topics from the top FREX terms. To prepare the topic prevalence summaries found in Table 3.2, I used a more extensive list of FREX terms for each topic than what is displayed in Figure 3.4. I also make use of the findThoughts function in stm to examine ads that are highly associated with the displayed topics.

Table 3.2 displays the content summaries of the topic prevalence contrast between far right and all other campaign appeals. In a large majority of countries, there is disproportionate emphasis on immigration issues on the far right, which is consistent with issue ownership. There are three notable trends in how the far right frames the immigration issue across Europe. First, many parties specifically emphasize Muslim migration and frame Islam as a unique threat to national values and cultural identity. Second, immigration is often tied to criminality as well as to women's safety. Third, it is often framed in terms of general Euroscepticism.

In a minority of cases, other parties integrate salient immigration issues into their existing ideological profiles. These include the Czech Republic, Denmark, Slovakia, and Sweden. For instance, in the Czech Republic, other parties do not reject the mandatory EU migrant quota scheme—but they do suggest that citizens should have a say in whether or not quotas are implemented. In Denmark and Sweden, there is an emphasis on imposing stricter border controls to prevent drug and weapons smuggling, as well as to curtail human trafficking. These are all more moderate appeals consistent with mainstream party ideologies. The exception is Slovakia, in which immigration appeals—mostly from the left-wing populist Smer-SD—are focused on implementing strict border controls to prevent the entry of migrants and refugees. In this case, a left-wing party uses ideologically incongruent immigration appeals to outbid the far right.

In three cases—Spain, Sweden, and Slovakia—I find no evidence of far right issue ownership on immigration. None of the most prevalent topics are immigration appeals. In Spain, VOX focuses on populist appeals contrasting the powerless masses (VOX patriots) and the powerful, globalist elites (leftist social democrats). In Sweden, the Sweden Democrats (SD) emphasizes the needs of rural Swedes and criticizes environmental restrictions for disproportionately disrupting rural livelihoods. In Slovakia, the far right focuses on welfare appeals and protecting the economic security of Slovak families. These exceptions demonstrate that the politics of the far right cannot simply be reduced to the politics of immigration; there are other social cleavages they may choose to emphasize. Shifting away from immigration issues is especially likely when other parties are actively 'out-bidding' on the immigration issue. This response to out-bidding is apparent in both Sweden and Slovakia.

Although this topic modeling application demonstrates heterogeneity in far right topic

[1]The remaining 10 stemmed FREX term plots are found in Appendix B.2.

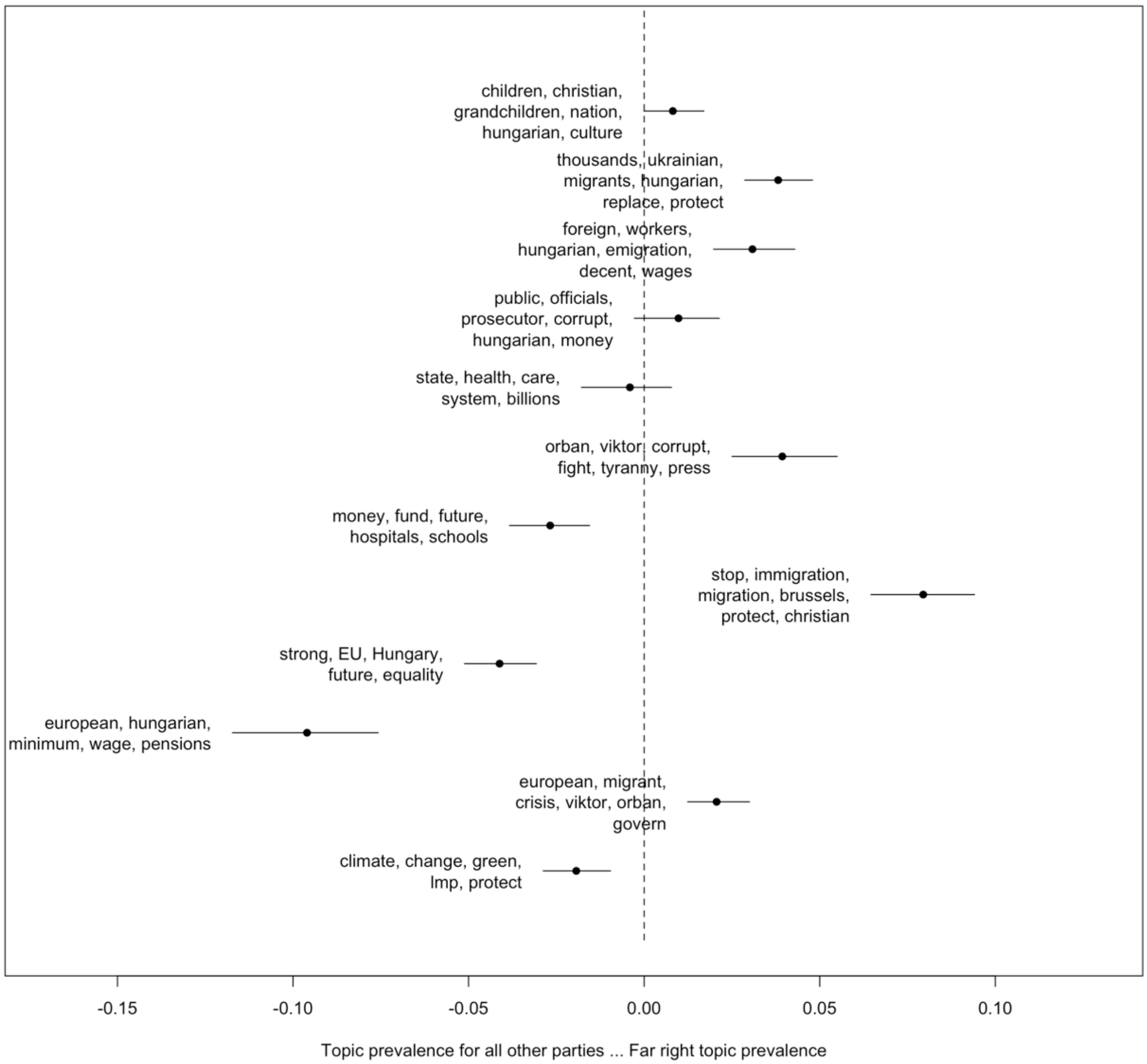


Figure 3.4: Sample Topical Prevalence Contrast (Hungary)

Sample plot of the change in topic proportions shifting from far right party documents (in this case, Jobbik & Fidesz) to other party documents (LMP, Momentum, DK, & MSZP). Hungary's far right has clear issue ownership over immigration-related topics. The top 'FREX' words are displayed, meaning words which are both frequent in and exclusive to a topic of interest.

Country	**Prevalent Topics, Other Parties**	**Prevalent Topics, Far Right**
Austria	EU climate & environmental protection standards; European citizenship; unions; agriculture & farmers	asylum policy; protection of external borders; Islamization & pro-immigration ideology in schools; income & corporate tax reform
Belgium	education; parental leave policy; affordable healthcare; taxes toward public transit; fair wages	Islam & illegal immigration; crime & border security
Czech Republic	senior housing & pensions; citizens' choice on migrant quotas; higher income & fair wages; technological development	issue of Muslim arsonists burning Christian churches; secure European borders; SPD patriotic movement; alliance with Orban & Salvini against Islam
Denmark	climate change & green transit; drug smuggling & border control; reform to cash assistance ceilings; investment in schools, social programs, & unions; aid programs for poor immigrants	migrants & border protection; Islam incompatible with Danish values; strict punishment for rape & other crimes; misuse of welfare funds to support foreigners; support for elderly care & pensions
Estonia	protect Estonian interests vis-a-vis Russia; cooperation on common EU defense, cyber, & climate policy	equality between EU members to ensure border controls on migration; defense of national interests in the EU; protection of farmers
Finland	cooperation on EU Arctic policy & on defense; climate change & reduction of global emissions; equal rights to work for women; healthcare & childcare support; prevention of tax evasion by the super-rich; regulation of hunting & fishing	protection of Finnish national interests in the EU; secure cross-border defense
Germany	education; environmental protections; jobs infrastructure & economic support; daycare & kindergarten fees; increase pensions & wages	Muslim violence against women; left-wing extremist antifa attacks on AfD; Syrian migration & crime; opposition to CDU on asylum policy; opposition to liberal gender ideology
Hungary	need to allocate more funds to hospitals & schools; increase pensions; strengthen relationship with the EU; climate & environmental protections	Christian culture for future generations; Ukrainian migrants stealing Hungarian jobs; Hungarian workers emigrating & reliance on cheap foreign labor; taking a stand against EU with border protections
Slovakia	fraud & corruption of former president Andrej Kiska; far right fascist threat; tax oligarchs & big business; support for healthcare; prosecute criminals & corrupt oligarchs; secure Hungarian & Austrian borders; education & teachers	support for healthcare, hospitals, & doctors; pension reform; welfare support for families & children
Spain	policy plans for Catalonia; revitalization of rural areas for youth; climate change; jobs in renewable energy; pension reform; economic security for freelancers & self-employed; gender wage gap	VOX as patriots & PSOE representing the elite & powerful; anti-globalism & anti-separatism; opposition to liberal gender ideology
Sweden	support LGBTQ, abortion, & human rights; environmental protections & climate change policy; prevent cross-border weapons & drug smuggling & human trafficking; cut EU funding to Hungary over rejection of EU migration policy; support EU-level migration policy	lower the price of gasoline; prioritize environmental policy needs of rural households; decrease fees & taxes on everyday resource consumption

Table 3.2: Content Summaries of the Topical Prevalence Contrast between Far Right and all Other Campaign Appeals

prevalence across Europe, the most prominent trend is far right issue ownership on immigration. In a majority of cases, mainstream parties have not appeased public pressure to 'confront' the issue. This strategic choice is less apparent from party manifestos, expert surveys, and media analyses. Party ad data, especially semi-private online ads, are well suited for analyzing issue ownership. With multiple election cycles of ads data, researchers will be able to examine the evolution of issue ownership across different party families and how these developments affected voter preferences over time.

Application 3. Is there an 'Immigration Void' in the Left-Right Policy Space?

Supply-side explanations of far right mobilization consider the political-institutional opportunity structures in which parties operate. Supply-side studies largely agree that the rise and persistence of the far right depends on the positioning of political parties within the left-right policy space (Kitschelt and McGann 1995). When mainstream parties occupy enough space on the ideological right, there is limited electoral opportunity for a far right challenger. However, when mainstream parties converge on more centrist positions, they leave a 'gap' in the electoral market. In this context, far right parties can quickly gain an electoral foothold and develop a reputation for being responsive on sensitive but salient issues (Carter 2005; Norris 2005; Arzheimer and Carter 2006).

In response to the emergence of a far right challenger, scholars expect mainstream parties to move further right on the challenger's signature issues. For example, numerous analyses based on party manifestos and expert surveys indicate that mainstream parties have gradually shifted to more extreme policy positions on immigration to win back voters (e.g., Bale et al. 2010; Van Spanje 2010; Rydgren 2012; Han 2015; Akkerman 2015). However, it is possible that a narrowed gap in formal party policy documents does not reflect a narrowed gap in informal party appeals and campaign rhetoric. Since the far right is incentivized to present a more moderate policy image in official documents, manifesto and survey data are insufficient to determine whether the mainstream right has shifted far enough to 'crowd out' the far right from voters' perspective.

This final comparative campaign ads data application answers the following question: are moderate parties' semi-private campaign appeals on immigration spatially proximate to the appeals of the far right? Through a scaling model of one-dimensional document positions called Wordfish, I examine whether mainstream parties have 'reclaimed the right' on the most salient issue to far right voters.

Wordfish is a word scaling algorithm developed by (Slapin and Proksch 2008) to estimate political positions from text documents. Like other text-based position estimates, this approach assumes that the relative word usage in a text document provides information about its placement along a spatial dimension. Wordfish builds a document-term matrix from text documents and assumes (1) that the distribution of words across texts follows a Poisson distribution, and (2) that this distribution can be modeled with document and feature fixed effects (Slapin and Proksch 2008). Wordfish is designed to evaluate text content on a single issue dimension, so I subset each party's ad corpus based on mentions of a set of immigration-related key words. Since Wordfish is an unsupervised scaling model, it only requires that researchers specify two texts which are roughly located at opposite ends of the latent dimension.

In this application, this means that the frequency with which a party mentions a specific word in a corpus of ads is drawn from a Poisson distribution, and the model treats each party's ad set as a separate party position. All party positions are estimated simultaneously, based on a manual specification of two parties' ad corpora that may be presumed as pro-immigration and anti-immigration in orientation. I defaulted to green parties as the pro-immigration and far right parties as the anti-immigration text samples unless prominent news coverage and expert reports from the country's election period indicated that the green party was moderate on immigration. In these rare cases, I substituted pro-immigration social democratic parties for the manual specification.

Using these model parameters, I evaluated all eleven countries on immigration policy and found three trends in party distribution. In a majority of countries, there is a sizeable spatial gap between the far right and all other parties. In six cases, there is at least a one standard deviation gap in the derived party positions. I interpret this gap as the absence of 'outbidding' on immigration in campaign appeals. In four of six cases—Germany, Spain, Estonia, and Finland—the spatial gap is closer to two standard deviations between the far right and its nearest ideological neighbor. More generally, the spatial distribution of parties in these six countries appears as expected; center-left green parties typically produce the most pro-immigration ad content—for example, Germany's Alliance 90/The Greens (GRÜNEN), Denmark's Red-Green Alliance (ENHL), and Austria's The Green Alternative (GREENS)—and other leftist parties are typically spatially close to the greens. Figures 3.5-3.6 illustrate these party position estimates.

In sum, a majority of cases indicate that informally stated party positions on immigration diverge to a greater extent than formally stated ones in party manifestos. Given this discrepancy, future qualitative research should examine two related questions: (1) are center-right party leaders and strategists aware of this large positional gap between far right campaigns and their own; and (2) if so, what incentives drive them to maintain moderate positions on immigration nonetheless. As the Facebook Ad Library accrues multiple election cycles of campaign data, it will present a unique opportunity to trace the conditions under which these positional 'appeal gaps' are narrowed.

There are other, less common spatial outcomes illustrated by the Wordfish algorithm. In Belgium and the Czech Republic, I find a much more even distribution between the far right and its nearest ideological neighbor on immigration. Figure 3.7 illustrates these party position estimates.

There are likely two factors driving these party distributions. First, both countries have structurally 'overcrowded' political fields; Belgium has eleven political parties and the Czech Republic has eight. The sheer number of political competitors leads to some crowding on issue positions. Second, this leads to a larger number of center-right parties that seek

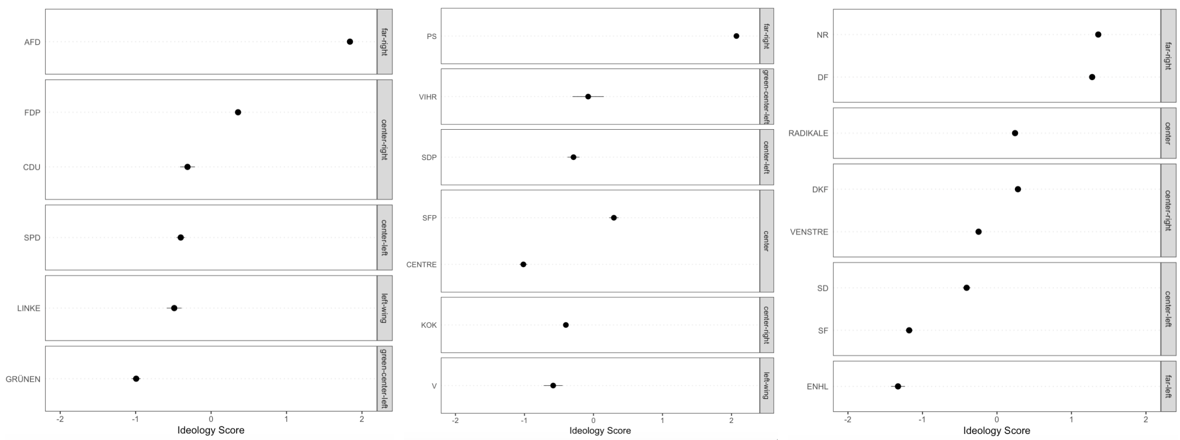


Figure 3.5: Estimated Party Positions on Immigration Policy, from left to right: Germany, Finland, Denmark

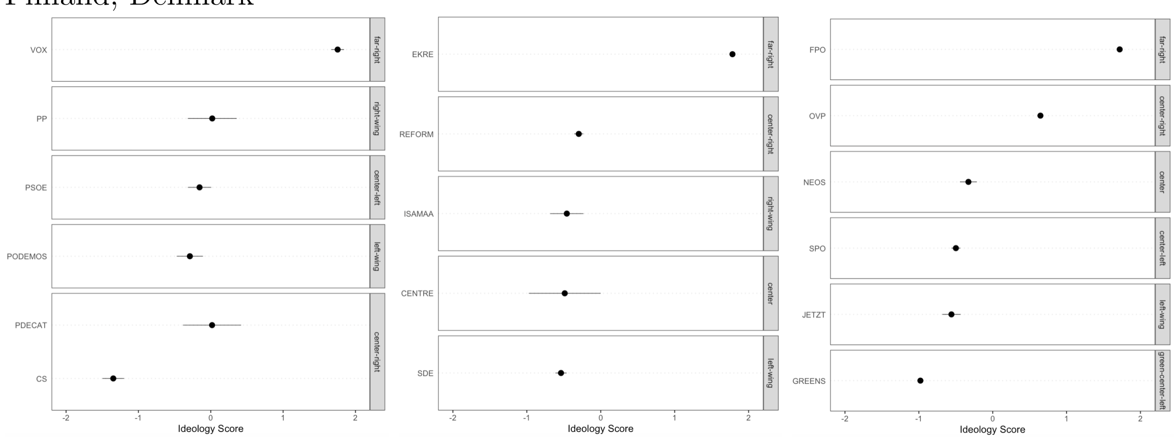


Figure 3.6: Estimated Party Positions on Immigration Policy, from left to right: Spain, Estonia, Austria

These countries exhibit the most uneven spatial distribution between the far right and their nearest ideological neighbors on immigration policy. Mainstream parties crowd around centrist and center-left positions on immigration. Note that in Denmark, there are 2 far right parties represented: the Danish People's Party (DF) and the New Right (NR). In all other countries, there is only one far right contender: the Alternative for Germany (AFD), the Finns Party (PS), Spain's Voice (VOX), the Conservative People's Party of Estonia (EKRE), and the Freedom Party of Austria (FPO). Each is located more than 1 standard deviation in spatial distance from its nearest ideological neighbor.

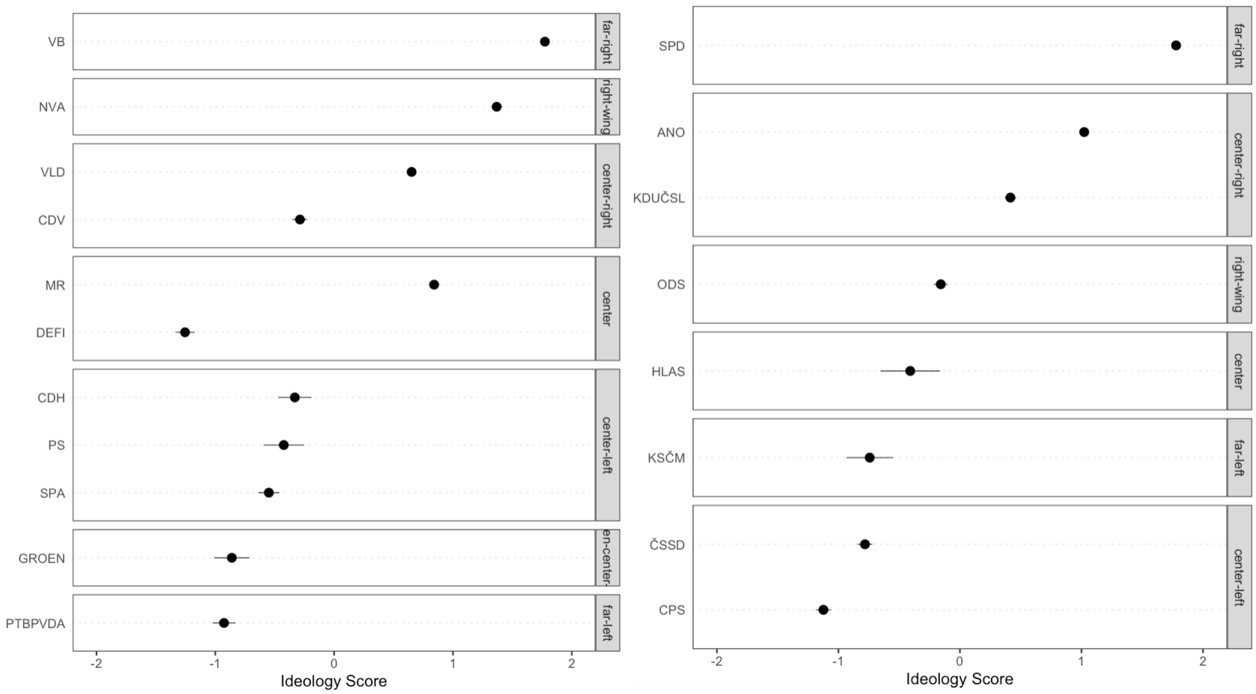


Figure 3.7: Estimated Party Positions on Immigration Policy, Belgium (left) and the Czech Republic (right)

These countries exhibit a more even spatial distribution between the far right (Vlaams Belang (VB) and Freedom and Direct Democracy (SPD), respectively) and their nearest ideological neighbors on immigration policy. Scholars theorize that under this spatial configuration, voters with strong anti-immigration preferences are less likely to perceive the far right as their only option.

to distinguish themselves from other center-right parties. If these are in fact important drivers, then the Belgian and Czech cases are not necessarily examples of outbidding the far right—but rather, of parties seeking to distinguish themselves in a crowded right-wing space.

Scholars theorize that within such a spatial configuration, voters with strong anti-immigration preferences are less likely to perceive the far right as their only option. Indeed, in both cases, the far right's closest ideological neighbor—the New Flemish Interest (NVA) and the Action of Dissatisfied Citizens (ANO)—currently hold the highest share of seats in parliament. With multiple election cycles of ad data, it will be possible to examine whether sustaining this spatial configuration facilitates the maintenance of power by center-right governments.

Finally, the Wordfish point estimates for Hungary, Slovakia, and Sweden indicate that it is possible for mainstream parties to outbid the far right on immigration. Notably, however, the ad data suggest something distinct from the manifesto data. While prior analyses of manifesto data suggest that mainstream contenders 'outbid' the far right by voicing support

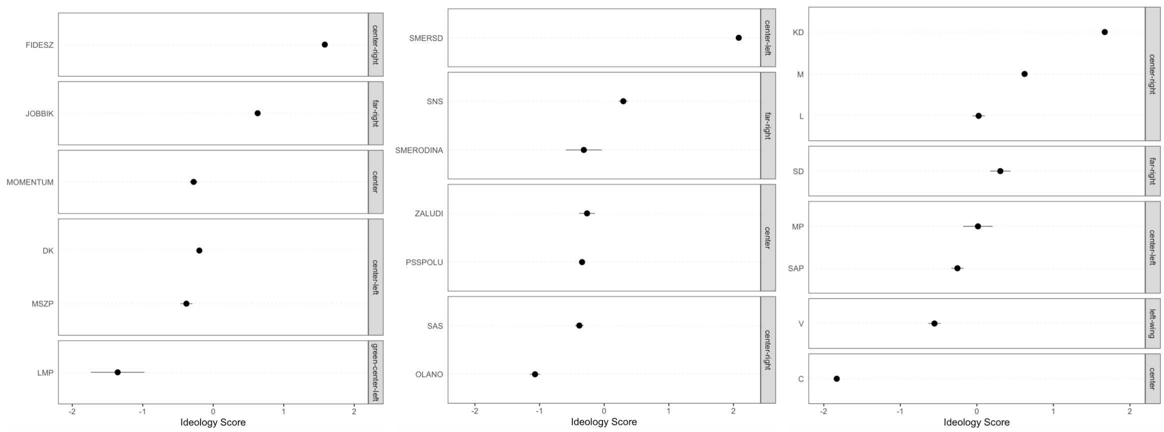


Figure 3.8: Estimated Party Positions on Immigration Policy, from left to right: Hungary, Slovakia, Sweden

These countries exhibit a pattern of outbidding the far right on immigration policy. In Hungary, the previously center-right Fidesz has moved dramatically right on immigration policy, forcing the previously far right Jobbik to move toward a more moderate position. In Slovakia, the center-left populist SMER-SD embraced radical anti-Muslim and anti-immigration discourse to try to stay in power. Although all Slovak parties are anti-immigration, SMER-SD's outbidding strategy makes the immigration appeals of other parties, including the far right Slovak National Party (SNS), appear centrist in relative terms. In Sweden, the center-right Christian Democrats (KD) and Moderate Party (M) outbid the far right Sweden Democrats (SD) by making more extreme appeals, while the SD has moved toward a more moderate position.

for similarly extreme policy positions(e.g., Bale et al. 2010; Van Spanje 2010; Rydgren 2012; Han 2015; Akkerman 2015), I find that these parties actually make more extreme appeals in semi-private campaign material. Figure 3.8 illustrates these party position estimates.

Could country experts have predicted the immigration ideology point estimates in Figure 3.8? In Hungary, they did: country experts have been raising the alarm for multiple years that the governing Fidesz coalition has taken a far right turn on immigration in an effort to maintain its supermajority in parliament vis-à-vis the Jobbik party (e.g., Bustikova and Guasti 2017). In Slovakia, the recent anti-immigration and anti-Muslim turn of the populist, center-left Smer-SD has been discussed by academics (e.g., Bustikova and Guasti 2017) as well as in policy briefs (e.g., Stiftung 2018). However, the extremity of Smer's campaign appeals compared to right-wing and far right Slovak parties was unexpected. In Sweden, while journalists have written about the far right Sweden Democrats's recent efforts at reputational moderation, no data indicate that either the center-right Christian Democrats (KD) or the center-right Moderate party (M) outbid them in anti-immigration appeals.

These spatial distributions illustrate the unique benefits of semi-private campaign appeals as text-as-data. Online ads demonstrate that both the far right and parties that seek to compete against them are incentivized to present controversial policy positions differently in public and in private. In most cases, this public-private gap implies an even larger distinction between the far right and all other parties on the issue of immigration.

3.6 Conclusion

The study of far right politics requires accurate data about party positions and appeals on these positions. These data are difficult to obtain, in part because far right parties face pressure to restrain the extremity of their policy positions. The gap between public and private ideological signals is potentially quite large, but scholars have overlooked it in part because informal party data is difficult to access. Although party manifestos, expert surveys, and media analyses have built up knowledge on some important dimensions, informal campaign messaging is a critical data source to further advance the field.

In the last year, social media data transparency initiatives have made it possible to explore the informal, microtargeted appeals parties use to mobilize voters. This article details why this new data source is so critical, both for the study of party strategy in general and for the study of far right politics in particular. It has numerous structural advantages over other data types. Namely, the Facebook Ad Library hosts perhaps the largest volume of text data on political ad campaigns—with more than 7.6 million ads available to date. Moreover, all political ad text is accompanied by important metadata, including (1) the time frame during which the ad was active on the platform, (2) audience location, and (3) basic audience demographics. This ad data is currently available for 36 countries across the globe.

The main objective of this article is to demonstrate the wide-ranging applications of these text data and illustrate how they can inform long-standing debates in the party politics literature. I evaluate three prominent theories on Europe's far right: (1) the notion of the far right as a grievance party; (2) the notion that the far right maintains issue ownership on immigration; and (3) the notion that far right parties succeed when there is an ideological gap on a salient issue dimension. I also demonstrate a range of computational tools that can be used to analyze ads data, including an unsupervised scaling model of document positions, structural topic models, and sentiment analysis.

Two of the main findings run counter to prior analyses of manifesto and expert survey data. First, party manifestos and expert surveys indicate that mainstream parties have gradually shifted to more extreme policy positions on immigration to win back voters (e.g., Bale et al. 2010; Van Spanje 2010; Rydgren 2012; Han 2015; Akkerman 2015). However, I find that this shift is not substantial enough to fill the ideological gap, because the far right uses its online ads to advocate for more extreme policy positions than what is stated in their manifestos. Second, while party manifestos indicate that mainstream parties integrate salient public issues into their ideological profiles to win back votes from the far right, I find that in a majority of countries, mainstream parties do not campaign heavily on immigration issues. I also confirm the theory that the far right is a distinctly grievance-based party. While previous work showed that demand for the far right stems from public grievance during national crisis, I demonstrate that they continue to ‘supply’ grievance-based narratives to their supporters once these crises have subsided.

There are certain limitations associated with online ads as data. Since advertisement online is voluntary, some parties opt out. Second, the Facebook Ad Library only hosts ads fielded since the end of 2018, so there is at most one election cycle of data for each country. Third, there is noise in online ads, because many are simple get-out-the-vote messages and not policy-oriented in content. While it is important to note these weaknesses, this article demonstrates that there is considerable potential for online ad libraries in party politics research. Examining sensitive political topics and extremism through semi-private ad data is especially critical. As this study demonstrates, the gap between public and private party information is often sizeable.

Chapter 4

A Machine Learning Approach to Detecting Polarization in Social Media Discourse

4.1 Abstract

Beyond social network analysis, political scientists have limited tools to measure polarization in online communities. We offer a new method that combines network analysis and supervised machine learning to detect changes in mass discourse on social media. We use our method to examine the effect of high salience news events on relevant political discussion on Twitter. We find that in day-to-day discourse, Twitter users across the ideological spectrum express diverse opinions on a range of political topics. Following news events salient to those topics, discourse becomes more extreme and partisan. Additionally, we find that the ideological tilt of discourse for users situated in centrist networks remains stable relative to users situated in more partisan networks. We demonstrate the potential of our supervised learning approach to address open questions in political science about who is polarized and when, and then discuss our method's limitations and practical challenges.

4.2 Introduction: Social Media's Role in Polarization

As the share of the global population that consumes political information through social media increases,[1] the technology once heralded as a tool of democratization has come under scrutiny. Numerous studies argue that social media exacerbate polarization and political information is unlikely to be transmitted if its content is ideologically cross-cutting, whereas

[1]For example, 66% of Brazilians, 62% of Portuguese, 58% of Spaniards, 51% of Swedes, and 51% of Americans surveyed by Reuters in 2017 reported receiving their news from social media (Newman et al. 2017).

hyper-partisan information spreads rapidly (Adamic and Glance 2005; Jamieson and Cappella 2008; Sunstein 2009; Yardi and Boyd 2010; Conover et al. 2011; Pariser 2011; Sunstein 2018). Other scholars remain more skeptical, arguing that social media facilitates exposure to messages from 'weak ties,' which can produce a moderating effect (Gentzkow and Shapiro 2011; Prior 2013; Barbera 2015b). This body of scholarship also suggests that real-world interactions provide fewer opportunities for cross-ideological exposure than digital ones (Gentzkow and Shapiro 2011).

The current literature relies on social network analysis (SNA) to measure the degree to which communities of social media users interact in segregated echo chambers and to draw inferences about polarization. While network analysis affords insights into the structural properties of digital communications, SNA overlooks what individuals within these networks actually say about politics. Although individuals' social networks correlate with their political positions and attitudes, it is possible that the opinions people hold within any given network vary widely. Moreover, with network analysis alone, it is not possible to distinguish between changes in users' political networks and changes in their political sentiment. As such, the inferential leap made using network analysis is often too large.

This paper offers a new approach to capturing polarization in online discourse, using a research design that narrows some of the current inferential gaps. Our data consist of thousands of randomly sampled tweets on partisan political topics (e.g., gun control and gun rights) proximate to relevant news events (e.g., a mass shooting). We first use SNA to create ideology scores for sampled users based on the network of political elites they follow on Twitter, and then use a machine learning approach to evaluate the content of their tweets. We aggregate these content evaluations by ideology score—comparing the speech of users with identical network scores pre- and post-event. This process allows us to evaluate changes in political discourse on social media following major news events, conditional on users' political networks. Since we use a repeated cross-section comparison, where different users comprise the samples before and after the event, holding their political networks constant and using the same search queries, we can attribute measurable differences in political content to social media news consumption during the event, rather than to the longer-term process of self-selection into media networks. These comparisons help us understand how social networks shape their members' political discourse during high news consumption periods.[2]

Our machine learning technique also improves on commonly used text analysis methods in political science. Most text-as-data studies rely on some combination of qualitative content analysis, sentiment analysis, and automated text analysis (using lexicon-based approaches). Since these methods have limited accuracy in classifying text, they are typically used to sort content into just two political categories (e.g., left or right; Remain or Brexit; hate speech or non-hate speech; radical Islamist or mainstream Islamist). Supervised machine learning provides a way to refine text classifications to be more granular, using labeled training data to distinguish between far right, moderate, and far left political text. We can use these

[2]These breaking news stories include the 2016 Brexit Referendum, the 2017 Las Vegas shooting, and Donald Trump's 2016 U.S. presidential election.

classifications to discern between two possible types of change in discourse: polarization and moderation. That is, we are able to discern between (1) the movement of discourse toward a political or ideological extreme, or (2) the movement of discourse toward a political or ideological center.

First, we briefly review the literature on social media and its role in polarization and discuss its main methodological limitations. Then, we describe our own research design and methods, detailing the advantages of a supervised learning approach to text analysis. Next, we present our findings and discuss how discourse shifts along the political spectrum as a response to salient news events. Finally, we discuss the conceptual and methodological limitations of our research and conclude by noting some important implications of network-driven mass polarization.

4.3 Current Approaches to Study Social Media and Polarization

A growing body of research focuses on micro-level indicators of social media-driven polarization. A small literature examines partisan Internet search behavior using web-administered behavior-tracking tools (Garrett 2009; Gentzkow and Shapiro 2011), online survey experiments (Munson and Resnick 2010), and laboratory-based web browsing experiments (Park et al. 2009). These studies find that people tend to consume opinion-reinforcing political information, but that they can be nudged toward politically cross-cutting information when they are (1) made aware of the partisan nature of their news consumption (Munson, Lee, and Resnick 2013, 425), or (2) when they are presented with a diverse set of news media options via news aggregators (Park et al. 2009). A larger literature turns to network analysis for insight. For example, Adamic and Glance (2005) demonstrate that partisan blogs link to each other more frequently and in a denser pattern compared to cross-ideological blogs. Conover et al. (2011) implement a network-based method of inference using cluster analysis and find that the Twittersphere is ideologically segregated in structure. Barbera (2015a) improves on this method using ideal-point estimation on a left-right ideological spectrum based on the media elites people follow on Twitter and finds that these ideal-point estimates map onto party registration records. His ideal-point estimates indicate that discourse on Twitter mostly occurs between those who share partisanship.

While these behavior-tracking and network-based methods predict general political orientation with good accuracy, they do not offer a substitute for text-based evaluation of political sentiment. Social media users with roughly equivalent ideological networks hold widely varied opinions on a range of salient political issues, as we will demonstrate below. Behavior-tracking and network-based approaches cannot capture such variation. An approach that combines SNA with text analysis helps bridge this analytical gap.

Text analysis is the processing of texts to extract structured information. Currently, researchers use a broad array of automated tools for text analysis. The literature, broadly speaking, utilizes machine learning, lexicon-based, rule-based, or statistical approaches. Machine learning combines multiple learning algorithms to detect text patterns through a training dataset. Lexical approaches assign polarity to text excerpts using the semantic orientation of words. Rule-based approaches classify information based on a set of human-coded terms that indicate opinions, feelings, and other subjective evaluations. Finally, statistical models identify latent patterns in text and use multinomial distributions, cluster analysis, and other methods to index it (Collomb et al. 2014).

These approaches to text analysis can lead to inaccurate (unreliable or biased) text classification, particularly when text is very brief. As Hopkins and King (2010) note, with short texts, the sparsity of the document-term matrix is one major source of trouble in developing accurate machine learning classifications (243). Overreliance on sparse terms within a text corpus weakens the generalizability of the results and leaves the model vulnerable to overfitting. Moreover, unsupervised machine learning methods are sensitive to topics rather than to sentiment expressed toward those topics, and may pick up on irrelevant features such as the informality of the author's writing style (Hopkins and King 2010, 231). Meanwhile, lexical and rule-based text analyses often rely on features such as social media hashtags for text classification, which is a problematic practice because hashtags are often used sarcastically or to inject content into ideologically opposed users' media feeds (Yardi and Boyd 2010; Conover et al. 2011). The lexicon-based approach proves particularly difficult to apply to social media data, because emoticons, colloquial expressions, abbreviations, and other text-level features possess semantic orientations but usually fall outside of the purview of opinion lexicons (Zhang and Liu 2011). These approaches frequently suffer from low recall (i.e., many false negatives), since they depend entirely on the presence of opinion words to determine the text's political orientation (Zhang and Liu 2011, 1).

Supervised machine learning with character n-gram representations offers a promising method to address these common pitfalls. Once a classifier is trained to determine the ideological placement of text, we map from pre-classified ideological text samples to unlabeled text to determine its ideological orientation. The next section will lay out our supervised machine learning method and address its benefits and limitations in the context of our research objectives.

4.4 Research Design: A Supervised Machine Learning Approach

First, we detail our event-based sampling procedure. Second, we discuss how we determined the partisanship of approximately 30,000 Twitter users with a Bayesian spatial model developed by Barbera (2015a), which leverages the political media they follow to calculate a network ideology score. Third, we discuss how we developed our labeled training dataset, which consists of more than 250,000 labeled tweets. Finally, we discuss the fastText machine learning algorithm, and how it was used to classify Twitter users' tweets as either far right, far left, or moderate in ideology.

Sampling Political Event Tweets

We used the full-archive Search Tweets API to sample thousands of tweets on different political topics chronologically preceding and following the timing of highly politicized news events. The same set of search parameters were used to collect the pre-and post event samples.[3] To carry out this sampling strategy, we had to define a range of events appropriate for the study of media influence. Our main objective was to choose highly politicized events, during the course of which political elites took distinct spins on how to interpret the event and its political implications. Each event had to be so widely discussed by the Twittersphere that we could broadly assume that thematic tweets posted immediately following the event were in response to it.[4] We performed a number of manual content checks on each event sample to corroborate this assumption.

For example, highly publicized mass shootings fit these requirements because they generate extreme media coverage and political 'spins' in addition to more mainstream left-leaning and right-leaning coverage. We chose the October 2017 Las Vegas shooting as one event and evaluated the Twitter conversation around the 'gun debate,' 'gun control,' and 'gun rights.' We sampled 5,000 tweets before and 5,000 tweets after each event using a combination of these three queries, using a five-day search window before and after the event. The day of the event was not sampled. The shooting was the deadliest in U.S. history, leaving 58 people dead and 851 injured. In the immediate aftermath of the shooting, more moderate media outlets focused on finding ways to implement certain 'common-sense' gun control restrictions whilst protecting gun rights. More extreme left-wing and right-wing media elites took darker, conspiratorial tones.[5] In addition, the far left proposed placing radical limits on gun ownership, while the far right used the event to stoke fear and push more gun sales.

The same procedure was used to sample tweets from two other news events: the June 2016 Brexit referendum and Donald Trump's November 2016 presidential election. In total, we evaluated the content of approximately 30,000 tweets across these three events. The geotagged locations of our sampled users indicate that tweets sampled from Brexit were predominantly U.K.-based, whereas the tweets sampled from the 2016 U.S. presidential election and the Las Vegas shooting were predominantly U.S.-based.[6] Given this geographic composition, our findings broadly speak to political media networks and their effects on mass opinion in the United States and the United Kingdom.

[3]These search parameters are detailed in Appendix C.1 Section A.

[4]For example, we would assume that tweets posted involving the phrases 'gun rights' or 'gun control' in the immediate aftermath of a high-profile mass shooting were written with the event in mind.

[5]For example, Infowars's extreme right-wing media figure, Alex Jones, falsely claimed that the Vegas shooter was an agent of the Islamic State, a leftist activist, and an anti-Trump radical.

[6]User location data is not available for the full sample, because many Twitter users opt out of sharing their location. However, geolocation is available for enough sampled users to make broad inferences about our study population.

Scoring Sampled Twitter Users' Ideological Networks

Next, we evaluated sampled users' political networks with data about the media elites they follow on Twitter. We used the tweetscores R package, which queries the Twitter REST API to identify a broad range of political accounts and media outlets users follow on Twitter. Then, using a Bayesian spatial model, tweetscores evaluates each user's network ideology based on the media they follow (Barbera 2015a).[7] We score sampled users as long as each follows at least one political Twitter account.[8] Once these media networks are mapped, the model uses correspondence analysis to project all users onto the latent ideological space and adjusts their ideological estimates to follow a normal distribution with a mean of zero and a standard deviation of one. Figure 4.1 illustrates the distribution of sampled tweeters pre- and post-event. In each event sample, the distribution is bimodal; ideologically extreme users are significantly more likely to 'chime in' to discuss political topics than moderates.

Developing a Training Dataset

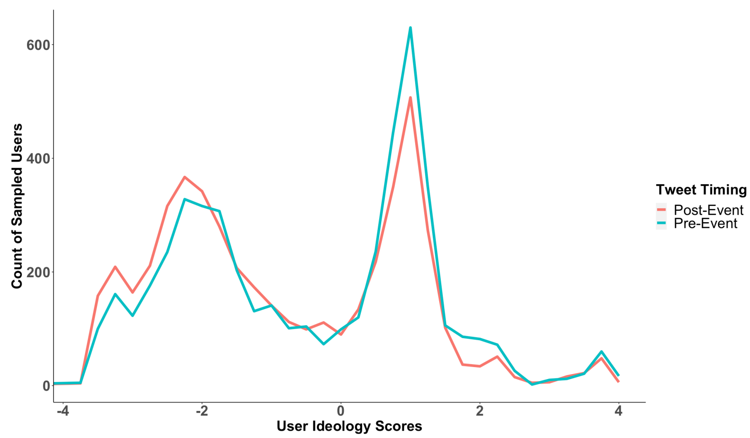
Supervised machine learning algorithms are difficult to implement because they involve an initial phase of manually labeling a large set of sample texts. However, manual in-house annotation poses sample size limitations, which can weaken the accuracy of complex learning models. We overcome this time-intensive manual classification phase by inferring the text classification of large corpora of tweets based on the media elites who wrote them, and labeling text based on this assumption. For example, when we query the tweets of far right conspiracy theorist Alex Jones, we assume his entire corpus of tweets fits within the 'far right' text classification, and automatically label each as far right in our training dataset.

Although this assumption introduces some error into the political classification of the unlabeled data, it is optimal for two reasons. First, since machine learning algorithms have to generalize from labeled training data to the unseen observations comprising the unlabeled data, it is most ideal to train data within its domain, so the model can "learn the unknown and underlying 'true' mapping that exists from inputs to outputs" (Brownlee 2018, 1). In other words, training in-domain improves the performance of the classifier because vocabulary and sentence structure differ substantially in unrelated domains (Silva et al. 2018, 224; Aue and Gamon 2005). Our pseudo-automated political classification method allows us to train a robust classifier on hundreds of thousands of tweets, bypassing the sample size limitations posed by manual annotation.

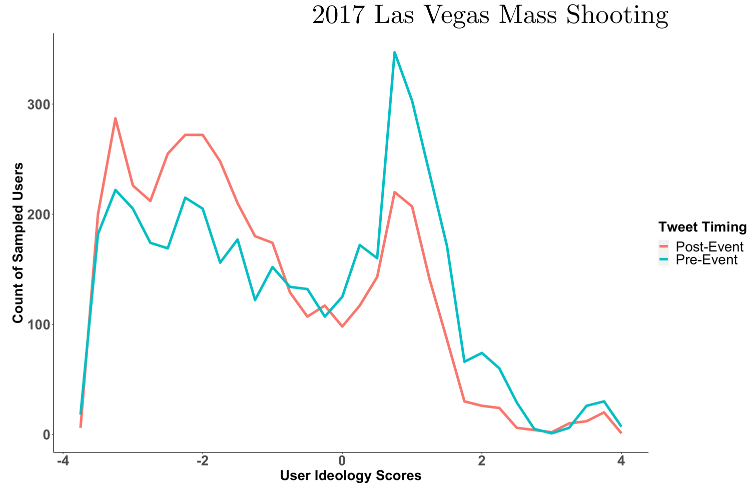
---

[7] We use the Metropolis-Hastings sampling algorithm to generate our ideology estimates. The package also contains an alternative specification—Maximum Likelihood estimation (MLE). As Barbera (2015a) points out, the MLE method is often preferable because of its speed and efficiency, since it is not sampling from the posterior distribution of the parameters. However, a major limitation of MLE is that it tends to give very narrow standard errors.

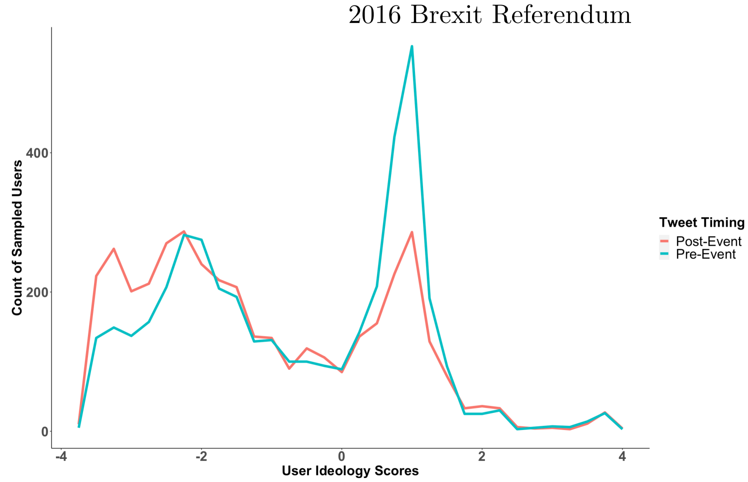
[8] We removed tweets where we could not estimate the network ideology of the user who tweeted it. Removal from the sample indicates that the Twitter user did not follow any type of political media or media elites on Twitter. Approximately three hundred sampled individuals could not be classified per sample of 5,000 tweets.



2017 Las Vegas Mass Shooting



2016 Brexit Referendum



2016 U.S. Presidential Election

Figure 4.1: Ideological Distribution of Sampled Twitter Users (Before & After Each Event)

Second, the manual classification of each tweet by coders may introduce significant error. Namely, it is difficult to classify brief online statements, whereas using our classification assumption, these statements are labeled in the context of a larger corpus and with knowledge of the speakers' political motivations. Our learning algorithm can detect patterns in political speech that human coders might be unable to recognize or label accurately. For example, given our assumption, when any media elite tweets about being an American, that tweet is automatically labeled in line with their political orientation, whereas a coder would likely determine that a tweet about being American has no political orientation. However, if in fact, far right elites discuss a supposedly neutral topic like being American at significantly higher rates than others, then the inclusion of these types of tweets in the training data provides useful information for the classifier, rather than noise. Similarly, the learning algorithm attenuates the influence of words and phrases it considers indiscriminate, based on their frequency distribution across political classifications.

Categorizing Tweets using fastText Machine Learning

Next, we evaluated the full corpus of tweets before and after each event using our training data. To construct this training set, we used the Twitter Premium API to sample approximately 250,000 tweets from media elites across the political spectrum. Our training data is comprised of American political elites and media outlets for the 2016 U.S. presidential election and the 2017 Las Vegas mass shooting and British media elites and outlets for the 2016 Brexit referendum.[9]

We have a measurement model where any given elite tweet is presumably a function of the elite's ideology plus an idiosyncratic personal term plus noise. Having a larger number of elites at any given point on the ideological spectrum reduces the idiosyncratic error, while having a large number of tweets per elite reduces the noise. To train the classifier on a large number of elites at each point on the spectrum, in addition to a large number of tweets per elite, we include at least 20 elites from each ideological camp (far left, moderate, far right). We sampled them with equal proportion, gathering 3,200 of each elite's most recent tweets. We labeled each elite's full set of tweets as one of three types of political sentiment—far right, moderate, or far left. This labeled set was used to 'train' our machine learning classifier, which then processed our unlabeled data (i.e., the 10,000 randomly sampled tweets per news event) and used information it had learned to label its contents. Using this process, the learning algorithm determined the political classification of each tweet in the event corpus.

This study employs an open-source machine learning approach called fastText. Facebook's AI Research lab developed the fastText machine learning library, which combines features of natural language processing with deep learning algorithms.[10] A number of fastText features make it highly efficient and relatively accurate at text classification tasks. We

[9]The training data parameters are detailed in Appendix C.1 Section C.

[10]See Joulin et al. (2016) for a detailed, technical overview. For a brief overview, see Mannes (2016)

briefly expand on these features and why they are suitable to the political classification of Twitter data.

First, fastText uses a hierarchical classifier (i.e., organized as 'trees') instead of a flat structure (i.e., lists), and represents texts as low dimensional vectors, which means that information learned about words in one category can be shared across all categories to improve text classifications (Grave 2016). By relying on neural networks, fastText learns low dimensional representations for all features in a text, and then averages these representations to make its classification predictions. The resulting models can then account for similarities between different features. For example, using low dimensional vectors, the model may be able to detect near-synonyms, such as 'woman' and 'female', and then treat these terms in a similar manner. This feature of fastText is particularly useful when classifying short texts, such as tweets, which had a 140-character limit during the sampled event periods.

Second, fastText utilizes character n-gram representations to create sub-word models. This means that the vector for a word is made of the sum of its character n-grams. For example the word vector "legal" is a sum of the vectors of the n-grams "", "leg", "lega", "legal", "legal_", "ega", "egal", "egal_", "gal", "gal_", and "al_".[11] This generates more neighbor context words for rarer and out of vocabulary words. It also addresses the sparse data problem of word-level n-grams and helps properly categorize approximate words (Kanaris, Kanaris, and Stamatatos 2006). These features are especially advantageous for classifying tweets, since it is common for social media data to contain spelling and grammatical errors, as well as the strange use of punctuation (Kanaris, Kanaris, and Stamatatos 2006, 3) and slang (Go, Bhayani, and Huang 2009).

A third advantage of fastText is that it utilizes a skipgram model, rather than a more commonly used continuous-bag-of-words (cbow) model. The main difference is that the skipgram model uses the training data to account for different probabilities of words in the corpus being a 'nearby word' and uses these neighbor associations to adjust its text classifications. When presented with classifying a target word within the unlabeled data, words that have higher probabilities of being 'nearby' are more heavily weighed in determining the target word's classification. In contrast, the cbow model predicts the classification of the target word according to its full context. The context is represented as a 'bag' of the words contained in a fixed size window around the target word (for example, 5 words on either side), and uses the sum of their vectors to make its prediction of the target. So, for example, the model receives more training samples of ('illegal', 'immigrants') than it does of ('illegal', 'Christians'). When the training is finished, if you input the word 'illegal', then a skipgram model will output a much higher probability for 'immigrants' or 'aliens' than it will for 'Christians.' Meanwhile, the cbow model takes all the words in a surrounding window, for example [true, Christians, oppose, illegal], and uses the sum of their vectors to predict the target.

In addition to these improvements over other learning algorithms, fastText allows for the

[11] This example assumes that the hyperparameters for the smallest n-gram is 3.

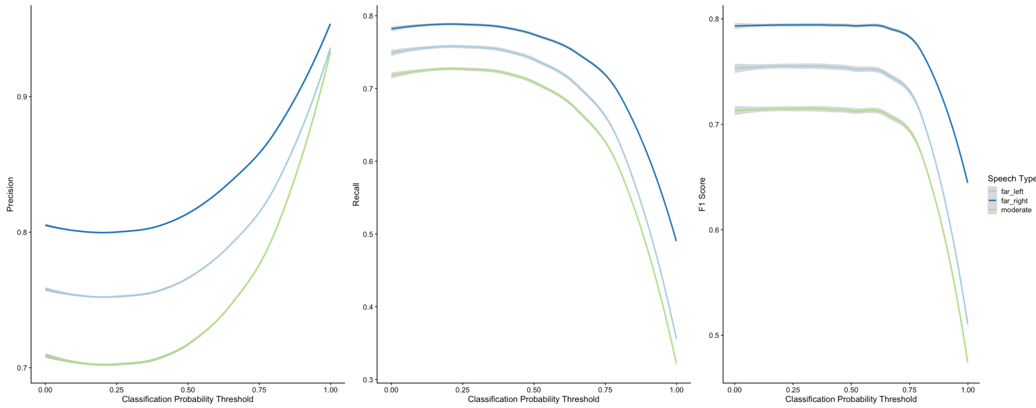
manual adjustment of other features to help optimize the model.[12] Figure 4.2 illustrates the precision, recall, and F1 score for each political label, by event. Precision refers to the number of labels the model correctly identified divided by the number of labels that it identified. For example, in sentiment analysis, if the model predicted there were 20 'positive' tone tweets, but only 16 were positive, the precision would be 80%. Recall refers to the total number of labels the model identified divided by the total number of labels possible. If there were actually 40 total 'positive' tone tweets in the above sentiment analysis, the recall rate would be 50%. By definition, this leads to a trade-off between precision and recall, since high precision aims to minimize false positives while high recall aims to minimize false negatives. We create a balanced classification model that optimizes for both in combination by using what is called the harmonic mean of precision and recall—the F1 score. For all three classification performance plots, the x-axis illustrates the precision cut-off when predicting text classifications. For example, if set to 0.9, the label will only be counted if the model predicts that the label is accurate with 90% confidence. Given that the F1 score remains stable between the 0.0 and 0.7 precision cut-off thresholds, we use a precision threshold of 0.7 and set the model to classify only tweets it can label with at least 70% confidence. At this precision cut-off, we are able to label a large majority of tweets with high confidence in the accuracy of the label.

Figure 3.2 also illustrates that, across all three events, fastText classifies far right speech with the highest precision and recall (ranging from 78-85%), far left speech with the second highest (ranging from 75-78%), and moderate speech with the lowest (ranging from 71-75%). Moderate speech is more difficult to classify because it is comprised of both moderate right-wing and moderate left-wing discourse. Additionally, the political extremes often feature distinct slang and terminology—making extreme speech easier for the algorithm to classify accurately .

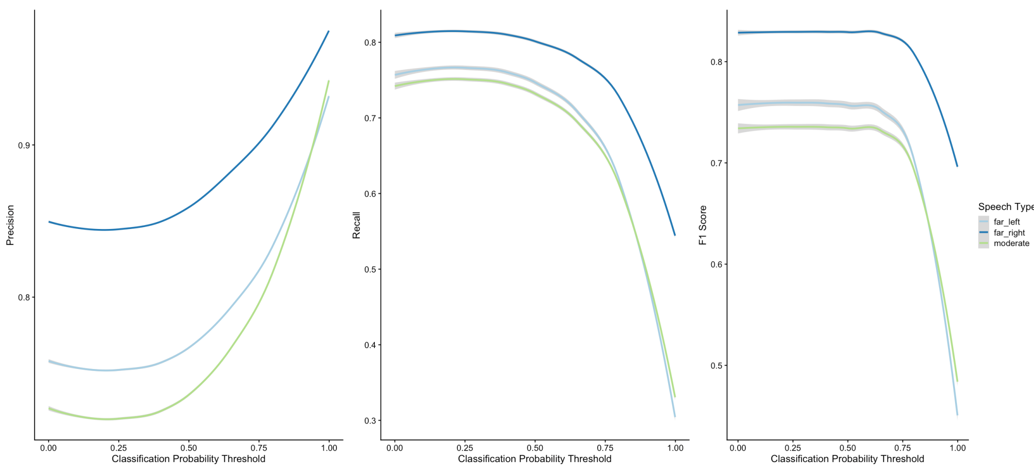
4.5 Findings: (A)symmetric Polarizations

The fastText algorithm produces three main findings. First, that in day-to-day discourse, social media users engage on political topics using relatively diverse speech that does not consistently align with their partisan networks. Second, following major news events, this variation in discourse is crowded out by more extreme, partisan content. Third, political centrists and partisans are affected differently by their political networks; moderates' discourse remains relatively stable, whereas partisans and extreme partisans become significantly more polarized. We run 100 bootstraps for each event to evaluate how consistent this finding is across iterations of our stochastic machine learning model and examine the results.

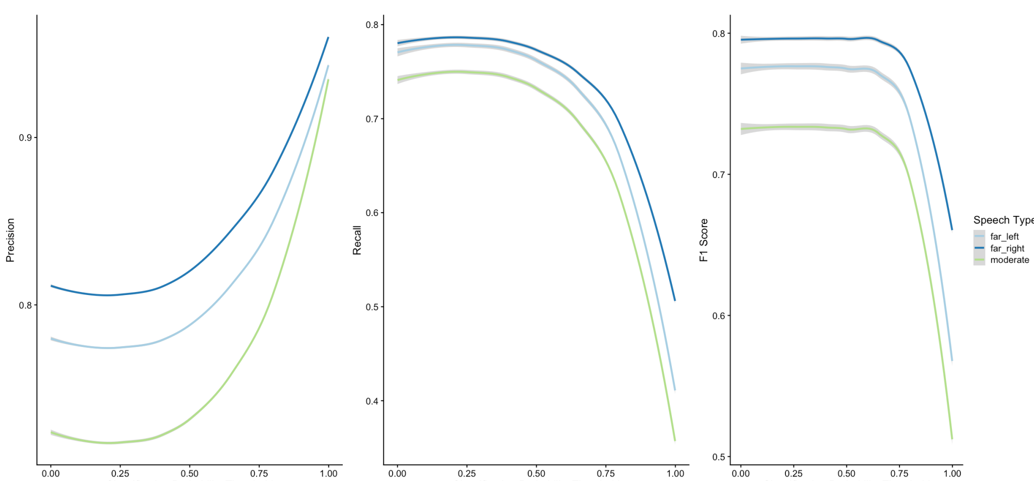
[12]A list of these features and our default settings for them can be found in the Appendix C.1 Section B.



2017 Las Vegas Mass Shooting



2016 Brexit Referendum



2016 U.S. Presidential Election

Figure 4.2: fastText Performance (Precision, Recall, F1 score), by Event

Measuring Ideological Network Effects

Figures 4.3, 4.4, and 4.5 summarize the results of 100 text classification bootstrap runs per event using the fastText machine learning algorithm. For each pre- and post-event sample, fastText used pre-labeled Twitter text (250,000 tweets from political media elites) to then label each sampled users' tweet. These labels are used to estimate the proportion of tweets along each point of the SNA ideological spectrum that can be classified as far left, moderate, or far right in political content.

Figures 4.3, 4.4, and 4.5 can be interpreted as follows. The x-axis represents the SNA ideological scores of each sampled Twitter user. Since Barbera's ideological estimates follow a normal distribution, 0 represents the ideological center; any user scored at or near zero follows centrist media elites, or a perfectly balanced combination of right-wing and left-wing elites. Scores to the left of center indicate that a user is ideologically left-leaning; they follow proportionally more left-wing media elites on Twitter than right-wing elites. Scores to the right of center indicate that a user is ideologically right-leaning; they follow proportionally more right-wing media elites on Twitter than left-wing elites. The farther left or right a user is along the ideological spectrum, the more ideologically skewed their network. At the far ends of the spectrum, users exist in increasingly extreme partisan media networks. For clearer visualization, users are grouped along the x-axis based on their network ideology scores in intervals of 0.25.

For each event, in the top figure, the y-axis represents the proportion of tweets that the fastText algorithm labeled as 'far left'; in the middle figure; the proportion of tweets labeled as 'moderate'; and in the bottom figure, the proportion of tweets labeled as 'far right' in ideological content. The observations are grouped into pre-event and post-event categories for comparison.

The pre-event content evaluations indicate that there is a fair amount of diversity in people's political attitudes (Lenz 2013; Broockman 2015). However, our data also show that there is a clear ideological bent to the tweets as predicted by who they follow. Sampled Twitter users in left-of-center networks tweet more ideologically left-leaning content, and those in right-of-center networks tweet more ideologically right-leaning content. Everyday political discourse on Twitter is clearly shaped by partisanship.

We use the pre-event and post-event differences in tweet classifications to understand how discourse changes along the SNA ideological spectrum. Following high-salience news cycles, we find that partisans become more polarized in their discourse. This polarization is sometimes symmetric and at other times asymmetric.

The 2017 Las Vegas mass shooting is an example of asymmetric left-wing polarization. Following the shooting, significantly fewer left-of-center users expressed far right sentiment, and significantly more expressed far left sentiment in their tweets. The shift toward far left discourse is evenly distributed for left-of-center users; the algorithm detects an increase of 5% in far left discourse following the mass shooting. Meanwhile, the tweets of users in centrist networks manifest the same ideological bend pre- and post-event. We also find evidence of moderation in discourse within right-of-center networks. The shift toward moderate discourse

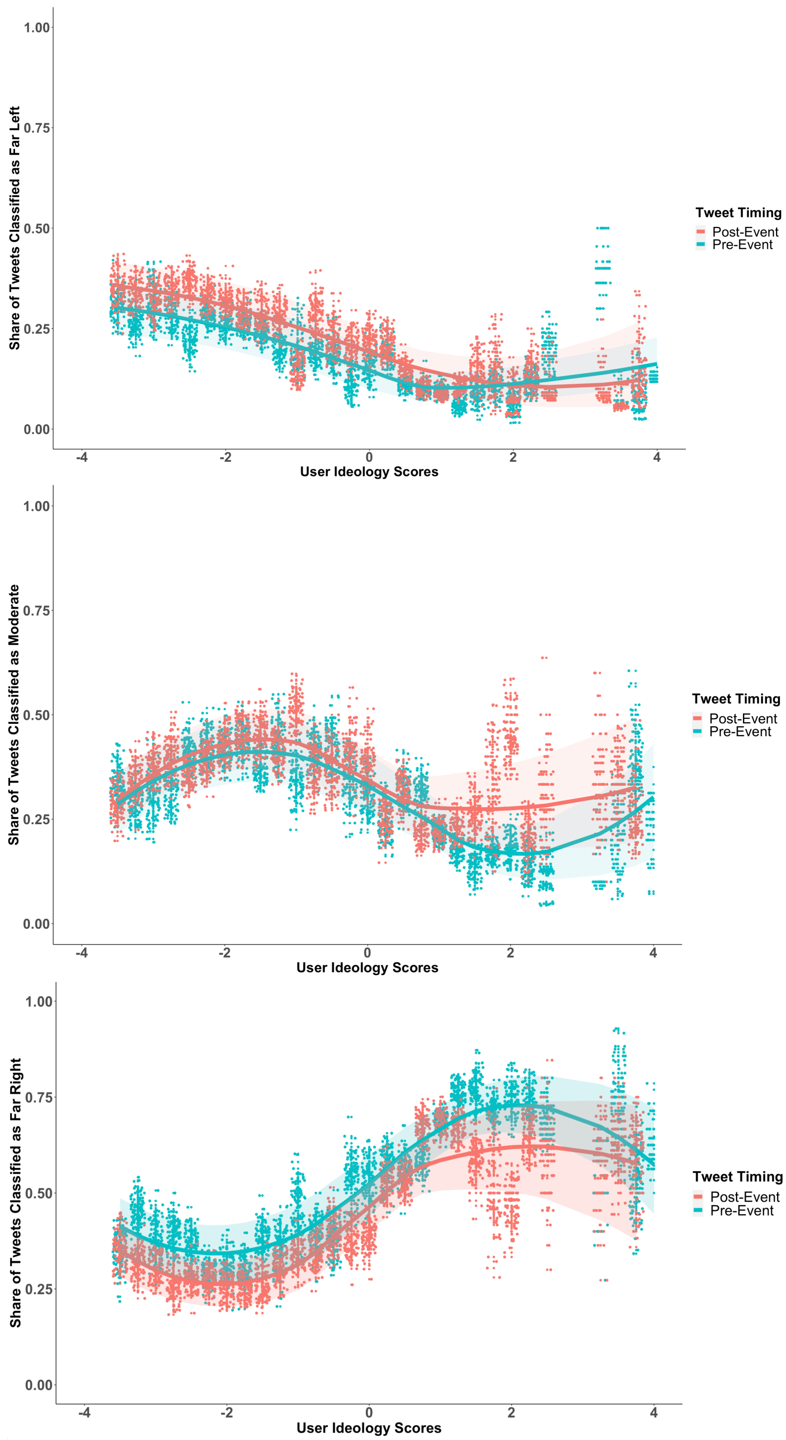


Figure 4.3: 2017 Las Vegas Mass Shooting Text Classifications along the SNA Ideological Spectrum (left to right) (100 Bootstraps, 95% Empirical Bootstrap Confidence Interval)

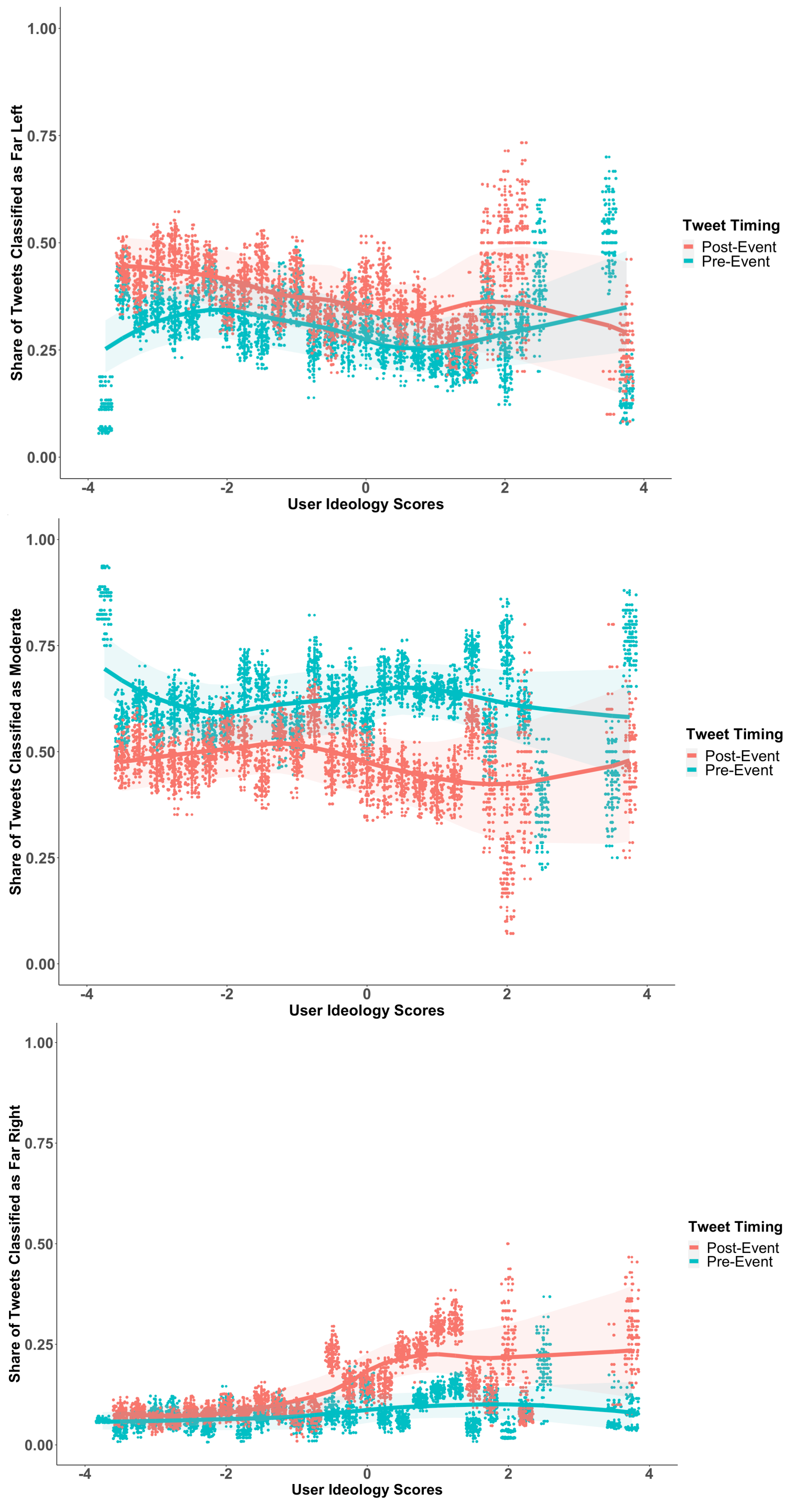


Figure 4.4: 2016 Brexit Referendum Text Classifications along the SNA Ideological Spectrum (left to right) (100 Bootstraps, 95% Empirical Bootstrap Confidence Interval)

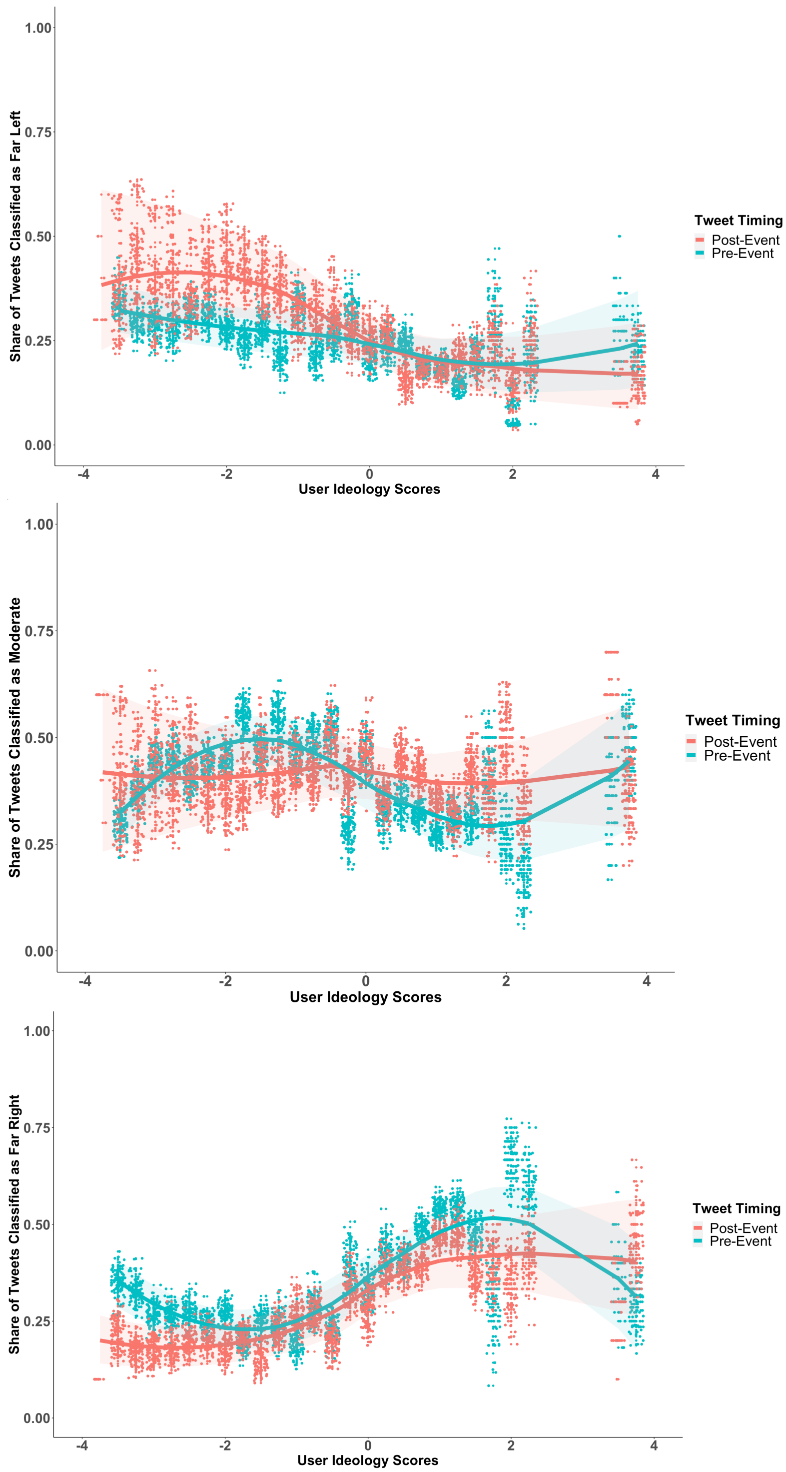


Figure 4.5: 2016 U.S. Presidential Election Text Classifications along the SNA Ideological Spectrum (left to right) (100 Bootstraps, 95% Empirical Bootstrap Confidence Interval)

is less apparent for the most extreme right-wing partisans.

The 2016 Brexit referendum is an example of symmetric polarization. Prior to the referendum, far right discourse was marginal; less than 10% of sampled discourse across the SNA ideological spectrum was classified as far right. Following the referendum, the share of far right discourse increased by 15% among right-wing partisans and by 5-10% among centrists. Left-of-center users exhibited no change in their share of far right discourse and a substantial increase in their share of far left discourse. The most extreme left-wing partisans increased their share of far left discourse by 15%. Prior to Brexit, more than 60% of speech across the SNA ideological spectrum was moderate in content. Moderate speech decreased substantially after the referendum—by more than 20% on the right and by more than 15% on the left. Our data indicate that the Brexit referendum marked an important moment of polarization in the country, after which 'pro-EU' and 'pro-Brexit' identities became more prominent in social media discourse. The change in discourse may also be indicative of political sorting into conservative and liberal camps based on the Brexit issue.

The 2016 U.S. presidential election is another example of asymmetric left-wing polarization. Following Donald Trump's election, fewer left-of-center Twitter users expressed far right sentiment, and significantly more expressed far left sentiment in their tweets. The increase in far left speech is smaller among moderate left partisans (5%), and stronger on the extreme left (10%). The tweets of users in centrist networks manifest the same ideological bend pre- and post-event. Although there is some moderation in the discourse of right-of-center users, the pre- and post-event estimates are statistically indistinguishable. We initially expected that Trump's election would mobilize the far right, but instead find that his election provided momentum for the emergence of more radical discourse on the political left.

Across these news events, we find that Twitter users embedded in partisan media networks exhibit diversity in their day-to-day political discourse on social media. On average, our learning model estimates that approximately one-fourth of Twitter users in left-of-center political environments express far right political sentiment, and vice versa, during baseline periods (e.g., 'pre-event'). Political discourse on a range of issues varies substantially between members of similar media networks. However, following major events, an increased share of discourse appears politically 'sorted' based on users' networks. Presumably, this is because information that users consumed in their media networks led them to align more closely with their own partisan camp. Social media users embedded in more moderate networks appear less affected by these political news cycles—perhaps because the elites and peers in their networks promote more balanced narratives.

4.6 Discussion and Conclusion

To conclude, we summarize the main advantages of our research design and methods. We then detail the limitations of our approach to spatial network analysis and of the fastText machine learning application. Finally, we discuss the substantive implications of our findings.

A New Method of Detecting Polarization in Social Media Discourse

Although findings from social network analysis suggest a strong relationship between social media consumption and political polarization, the evidence to date is limited because polarization is inferred based on the structural characteristics of social networks, rather than based on users' expressed beliefs. The current literature is further limited by the inability to properly parse causes (e.g., social media news consumption radicalizing users) from effects (e.g., extreme partisans opting into homogeneous social media environments). Through our 'salient events' research design, we study the effects of social media consumption on a short time horizon and avoid capturing the effects of the longer-term process of self-selection into different networks.

We improve on current methods of detecting polarization in four ways. First, we combine traditional social network analysis with a machine learning approach to text analysis to examine political communication across the ideological spectrum of Twitter users. By combining these methods, we reduce an inferential leap made by prior studies on polarization. Specifically, we narrow the gap between what social media users' networks indicate about their partisanship and what they actually say about politics. Second, we employ machine learning tools rather than lexicon-based methods of automated text analysis to classify political text. This improves recall and avoids problems inherent in lexicon-based methods—namely, their inability to accurately identify and label coded speech, trolling, and sarcasm, among other text features. Third, we train our classifiers using three labels, rather than sorting political text into binary 'conservative-leaning' or 'liberal-leaning' categories. This allows us to examine political polarization with greater granularity than prior research. It also helps explain what happens to moderate sentiment during salient political events. Fourth, we examine multiple political events across two English-speaking countries. This increases the generalizability of our findings.

Notably, we find that salient events have similar effects across the British and American political systems. In both, these events lead to polarization in political discourse. However, moderate discourse is much more common in British social networks; our algorithm classifies a large majority of content across the SNA ideological spectrum as moderate prior to the Brexit referendum. In American media networks, moderate political speech accounts for just 30-50% of all sampled political discussion on contentious issues, indicating that the U.S. information ecosystem is significantly more polarized. Although there is insufficient comparative research on social media and polarization to confirm or refute this observation, Barbera (2015b) finds that exposure to "high political diversity" is substantially more common for German and Spanish Twitter users than it is for U.S. Twitter users (see Barbera 2015b, 18). This finding is consistent with ours.

The Challenges of Categorizing Political Tweets

Our methods have a number of limitations. First, we use Barbera's tweetscores package to score sampled Twitter users and determine the political slant of their social networks. The tweetscores package uses a Bayesian spatial model to evaluate all Twitter users along a one-dimensional left-right political scale, based on the pre-estimated scores of hundreds of prominent media elites. Barbera notes that his model performs best in the United States, where economic and social positions are highly correlated (Barbera 2015a, 17). Since economic and social positions are multidimensional in European countries, our reliance on Barbera's latent spatial model becomes more problematic in the case of Brexit, potentially weakening the validity of our findings for that event. Future comparative studies of social media should develop more reliable point estimates that account for separate social and economic spatial dimensions. Moreover, insofar that we conceptualize partisan polarization as one-dimensional phenomena, political scientists must create more complex 'tests' for its presence or absence outside of the United States.

Second, although fastText is a highly efficient machine learning model on par with deep learning classifiers in terms of accuracy, it has weaknesses. Since fastText gains efficiency through hierarchical classifiers rather than through flat, list-like structures, bias propagates. For example, if an n-gram is wrongfully categorized in a 'parent' node of the hierarchical 'tree' it will inevitably be misclassified in the 'child' nodes branching off of that tree as well. Errors in classification can compound, with limited recourse to identify and correct these biases. Moreover, because fastText leverages character n-grams, the model requires incredibly large datasets to learn generalizable embeddings (Major, Surkis, and Aphinyanaphongs 2017, 3). Although we leverage a large corpus of training data, we still face a trade-off in precision.

Our third weakness is conceptual. Our training data relies on the assumption that the texts of each media elite can be labeled consistently as far right, moderate, or far left in content. Machine learning researchers who typically use full manual annotation to create training sets may be surprised at the accuracy of a pseudo-automated training set, but we performed hundreds of qualitative checks to ensure that this assumption generally holds. It is likely the case that media elites' tweets can be classified consistently because the current media ecosystem is highly fragmented, and political elites are incentivized to cater to their narrow segment of supporters. The net benefit of pseudo-automated classification is that the fastText learning model requires incredibly large datasets to learn generalizable embeddings, and our method for creating a training set produces a corpus featuring more than 250,000 labeled tweets and approximately 7 million words.

The Implications of Rapid Partisan Polarization

Our findings are consistent with political scientists who argue that polarization is occurring through an erosion of heterogeneous political positions within different political communities, both elites (Poole and Rosenthal 1984; Adams 1997; Fiorina and Levendusky 2006) and masses (Fiorina and Abrams 2008; Hetherington 2009; Mason 2015). Our addition to this literature is to empirically demonstrate how salient events facilitate the process of mass partisan polarization. Although the current literature conceptualizes partisan polarization as a multi-decade phenomenon, we provide data across multiple cases that suggest that it can occur on a short time horizon, as a result of shocks to the landscape of political news and public discourse. In fragmented social media environments, low attention individuals absorb and interpret political information through a highly stylized political lens. This is an important supplement to the body of work on long-term, generational trends on polarization.

One future extension of our machine learning approach is to measure the rate of decay in rapid polarization. We currently only leverage a five day window pre- and post-event to quantify the immediate effects of news consumption during contentious news periods. However, the implications of a one or two week decay in polarized discourse are different from the implications of a slower rate of decay. In addition, it is important to examine whether discourse ever recedes to 'baseline' levels of polarization.

Other Applications of our Machine Learning Approach

Political scientists can gain leverage on important behavioral questions through our research design. In particular, machine learning approaches can help refine our understanding of the complex relationships between public opinion, polarization, and social media consumption, given their wide range of applications compared to lexicon-based methods or sentiment analysis. More broadly, machine learning text classification is a necessary supplement to the wealth of social network analyses in the literature.

Relatedly, we urge researchers to use our machine learning approach to adopt more fine-grained political text classifications. When studies use binary classifications of 'left-wing' and 'right-wing' political speech, they fail to disaggregate subgroups with distinct political discourse and behaviors. Studies working with binary text classifications also de facto disregard moderate political discourse. As our machine learning application demonstrates, a large range of political discourse online is moderate, and the design choice to include moderate speech as a category has important analytical consequences.

Conceptually, we believe our 'salient news shock' framework can be applied to study other important questions about social media and help begin to parse cause and effect online. Current approaches to social media analysis are sufficient for descriptive research, but more complex frameworks are necessary to increase the number of causal studies on social media.

Lastly, comparative social media research is necessary to understand the institutional and social mechanisms that exacerbate polarization. We believe that our machine learning framework is suitable for such research. Our training data was comprised of relatively well-known and country-specific media elites. With the help of country-specific media experts, developing reliable training data is relatively simple for many countries. In short, our machine learning approach offers a flexible path forward for comparative studies of social media environments.

Chapter 5

Conclusion

This dissertation detailed three novel ways to conceptualize and measure the distribution of extreme political preferences in democratic electorates. My findings have important implications for the study of far right extremism in particular. Scholars have pointed to various explanations for the rise of far right extremism, but few can explain its persistence and continued growth. My dissertation fills this gap by empirically distinguishing between staunch and contingent extremists—groups that are mobilized at different temporal points through different means.

The main chapters contribute a few other notable findings for further consideration. Chapter 2 leveraged two empirical analyses that show how the perceived popularity of the far right activates voters, especially in places where the actual party popularity is low. These analyses produce a striking finding, namely that in the absence of actual popularity, perceived popularity can be powerful in mobilizing extremists. The extent to which perceptions of party popularity are malleable necessitates further study. The current political information environments in the United States and in parts of Europe are heavily fragmented and feature biased reporting of political data. In these contexts, the findings of Chapter 2 are especially worrisome. Future research should examine (1) how political extremists come to believe biased information; (2) how far information can be from the truth and still be perceived as credible; and (3) what incentive structures promote skepticism of attitude congruent but false information.

In Chapter 3, I applied sentiment analysis to political ads and concluded that far right parties make grievance appeals at significantly higher rates than other parties. Moreover, their ad content features more heavily in two emotive appeal types: fear and disgust. There is a vast experimental literature in political psychology that suggests that these emotive appeals are effective as mobilization tools. However, no studies to date evaluate the efficacy of these appeal types in online ads. Ad experiments that vary emotive appeals and evaluate ad impressions, click-through-rates, and other measures of interaction would serve as an important extension on Chapter 3. It is possible that digital ad experiments replicate the findings of more traditional lab experiments, but emotive appeals online may also have distinct effects.

In Chapter 4, I applied a supervised learning algorithm to social media discourse and detected either symmetric or asymmetric polarization following major news events. This suggests that polarization can occur rapidly, and that major political events are vulnerability points for political radicalization. My finding stands in stark contrast to the current literature, which conceptualizes partisan polarization as a slow phenomenon that evolves over multiple decades. More rigorous follow-up research is necessary to understand the decay rate of polarization in social media discourse. Future research should also examine whether any permanent effects accumulate over one or more similar major news events. Two examples of major news events that frequently repeat, inundate social media discourse, and have the potential to produce 'sticky' polarization effects over time are terror attacks and mass shootings. My machine learning method can be used to shed light on these critical questions about the durability of polarization.

To conclude, the research approaches introduced in this dissertation are easily replicable and widely applicable. The future avenues for research I highlighted touch on a number of literatures. Most importantly, these methods help bridge the gaps between public and private political extremisms.

Bibliography

Adamic, L. A. and N. Glance (2005). "The Political Blogosphere and the 2004 US Election: Divided they Blog". In: *Proceedings of the 3rd International Workshop on Link Discovery (ACM)*, pp. 36–43.

Adams, G. D. (1997). "Abortion: Evidence of an Issue Evolution". In: *American Journal of Political Science* 41.3, pp. 718–737.

Adams, J. and Z. Somer-Topcu (2009). "Policy adjustment by parties in response to rival parties' policy shifts: Spatial theory and the dynamics of party competition in twenty-five post-war democracies". In: *British Journal of Political Science* 39.4, pp. 825–846.

Akkerman, T. (2015). "Immigration policy and electoral competition in Western Europe: A fine-grained analysis of party positions over the past two decades". In: *Party Politics* 21.1, pp. 54–67.

Albertson, B. and S. K. Gadarian (2015). *Anxious politics: Democratic citizenship in a threatening world*. New York, NY: Cambridge University Press.

Anderson, E. and N. Vinocur (2018). *Marine Le Pen proposes renaming party National Rally*. URL: `https://www.politico.eu/article/marine-le-pen-french-far-right-proposes-renaming-national-front-party-national-rally/`.

Arceneaux, K., M. Johnson, and C. Murphy (2012). "Polarized political communication, oppositional media hostility, and selective exposure". In: *The Journal of Politics* 74.1, pp. 174–186.

Arzheimer, K. and E. Carter (2006). "Political opportunity structures and right-wing extremist party success". In: *European Journal of Political Research* 45.3, pp. 419–443.

Aue, A. and M. Gamon (2005). "Customizing Sentiment Classifiers to New Domains: A Case Study". In: *Proceedings of Recent Advances in Natural Language Processing (RANLP)*.

Bakker, R. et al. (2014). "Anchoring the experts: Using vignettes to compare party ideology across countries". In: *Research and Politics* 1.3.

Bale, T. et al. (2010). "If you can't beat them, join them? Explaining social democratic responses to the challenge from the populist radical right in Western Europe". In: *Political studies* 58.3, pp. 410–426.

Barbera, P. (2015a). "Birds of the Same Feather Tweet Together: Bayesian Ideal Point Estimation using Twitter Data". In: *Political Analysis* 23.1, pp. 76–91.

— (2015b). "How Social Media Reduces Mass Political Polarization. Evidence from Germany, Spain, and the US". In: *Presented at the 2015 American Political Science Associ-*

ation Conference. URL: http://pablobarbera.com/static/barbera_polarization_APSA.pdf.

Bartels, L. M. (1988). *Presidential primaries and the dynamics of public choice*. Princeton University Press.

Bartolini, S. and P Mair (1990). "Policy competition, spatial distance and electoral instability". In: *Political Analysis* 13.4, pp. 1–16.

Baumann, M. and M. Gross (2016). "Where is my party? Introducing new data sets on ideological cohesion and ambiguity of party positions in media coverage". In: *Arbeitspapiere/Mannheimer Zentrum für Europäische Sozialforschung= Working papers* 167.

Benkler, Y., R. Faris, and H. Roberts (2018). *Network propaganda: Manipulation, disinformation, and radicalization in American politics*. Oxford University Press.

Benoit, K. and M. Laver (2006). *Party policy in modern democracies*. Routledge.

Boas, T. C., D.P. Christenson, and D.M. Glick (2018). "Recruiting large online samples in the United States and India: Facebook, Mechanical Turk, and Qualtrics". In: *Political Science Research and Methods*, pp. 1–19.

Bonikowski, B. (2017). "Ethno-nationalist populism and the mobilization of collective resentment". In: *The British Journal of Sociology* 68.1, pp. 181–213.

Bos, L., W. Van der Brug, and C. De Vreese (2010). "Media coverage of right-wing populist leaders". In:

Brand, J.J. (1999). "Development, implementation and evaluation of multiple imputation strategies for the statistical analysis of incomplete data sets". In:

Broockman, D.E. (2015). "Public Opinion and Political Representation". In: *Doctoral dissertation*. URL: https://escholarship.org/content/qt0vg8049w/qt0vg8049w.pdf.

Brownlee, J. (2018). "Why Do Machine Learning Algorithms Work on Data That They Have Not Seen Before?" In: URL: https://machinelearningmastery.com/what-is-generalization-in-machine-learning/.

Bustikova, L. and P. Guasti (2017). "The Illiberal Turn or Swerve in Central Europe?" In: *Politics and Governance* 5.4, pp. 166–176.

Bustikova, L. and H. Kitschelt (2009). "The radical right in post-communist Europe. Comparative perspectives on legacies and party competition". In: *Communist and Post-Communist Studies* 42.4, pp. 459–483.

Buuren, S. V. and K. Groothuis-Oudshoorn (2010). "mice: Multivariate imputation by chained equations in R". In: *Journal of statistical software* 4, pp. 1–68.

Carter, E. (2005). *The Radical Right in Western Europe: Success or Failure*. Manchester, New York: Manchester University Press, and Palgrave.

Chaiken, S. (1987). "The heuristic model of persuasion". In: *Social influence: the ontario symposium* 5, pp. 3–39.

Cole, A. (2005). "Old right or new right? The ideological positioning of parties of the far right". In: *European Journal of Political Research* 44.2, pp. 203–230.

Collomb, A. et al. (2014). "A Study and Comparison of Sentiment Analysis Methods for Reputation Evaluation". In: *Rapport de recherche RR-LIRIS-2014-002*.

Conover, M. et al. (2011). "Political Polarization on Twitter". In: *Fifth International AAAI Conference on Weblogs and Social Media (ICWSM)* 133, pp. 89–96.

Curini, L. (2010). "Experts' political preferences and their impact on ideological bias: An unfolding analysis based on a Benoit-Laver expert survey". In: *Party Politics* 16.3, pp. 299–321.

De Vries, C. E. and G. Marks (2012). "The struggle over dimensionality: A note on theory and empirics". In: *European Union Politics* 13.2, pp. 185–193.

Debus, M., Z. Somer-Topcu, and M. Tavits (2016). *Comparative Campaign Dynamics Dataset*.

Dennison, J. and M. Goodwin (2015). "Immigration, issue ownership and the rise of UKIP". In: *Parliamentary Affairs* 68.1, pp. 168–187.

Department, Statista Research (2019). *Total number of Facebook users in France in 2018, by age group (in millions)*. URL: `https://www.statista.com/statistics/866813/french-use-internet-facebook/`.

Dinas, E. and K. Gemenis (2010). "Measuring parties' ideological positions with manifesto data: A critical evaluation of the competing methods". In: *Party politics* 16.4, pp. 427–450.

Elmelund-Præstekær, C. and H. Mølgaard-Svensson (2014). "Negative and personalized campaign rhetoric: Party communication and media coverage of Danish parliamentary elections 1994–2011". In: *World Political Science* 10.2, pp. 365–384.

Ennser, L. (2012). "The homogeneity of West European party families: The radical right in comparative perspective". In: *Party Politics* 18.2, pp. 151–171.

Eurostat (2017). *Are you using social networks?* URL: `https://ec.europa.eu/eurostat/web/products-eurostat-news/-/DDN-20170713-1`.

Fiorina, M. P. and S.J. Abrams (2008). "Political Polarization in the American Public". In: *Annual Review of Political Science* 11, pp. 563–88.

Fiorina, M. P. and M.S. Levendusky (2006). "Disconnected: The Political Class versus the People". In: *Red and Blue Nation? Characteristics and Causes of America's Polarized Politics*. Ed. by S. N. Pietro and D.W. Brady. Washington, DC: Brookings Institution Press, pp. 49–71.

Gabel, M. J. and J. D. Huber (2000). "Putting parties in their place: Inferring party left-right ideological positions from party manifestos data". In: *American Journal of Political Science*, pp. 94–103.

Garrett, R. K (2009). "Echo Chambers Online?: Politically Motivated Selective Exposure Among Internet News Users". In: *Journal of Computer-Mediated Communication* 14.2, pp. 265–285.

Gattinara, P. C. and C. Froio (2019). "Getting 'right'into the news: grassroots far-right mobilization and media coverage in Italy and France". In: *Comparative European Politics* 17.5, pp. 738–758.

Gattinara, P. C. and L. Morales (2017). "The politicization and securitization of migration in Western Europe: Public opinion, political parties and the immigration issue". In: *In Handbook on Migration and Security*.

Geer, J. G. (1988). “The effects of presidential debates on the electorate’s preferences for candidates”. In: *American Politics Quarterly* 16.4, pp. 486–501.

— (1989). “Nominating presidents: An evaluation of voters and primaries”. In: *Greenwood Pub Group* 286.

Gentzkow, M. and J.M. Shapiro (2011). “Ideological Segregation Online and Offline”. In: *The Quarterly Journal of Economics* 126.4, pp. 1799–1839.

Giugni, M. et al. (2005). “Institutional and discursive opportunities for extreme-right mobilization in five countries”. In: *Mobilization: An International Quarterly* 10.1, pp. 145–162.

Go, A., R. Bhayani, and L. Huang (2009). “Twitter Sentiment Classification using Distant Supervision”. In: *Technical report*.

Grave, E (2016). “Faster, Better Text Classification!” In: URL: `https://github.com/facebookresearch/fastText/blob/master/website/blog/2016-08-18-blog-post.md`.

Grimmer, J. and B. M. Stewart (2013). “Text as data: The promise and pitfalls of automatic content analysis methods for political texts.” In: *Political analysis* 21.3, pp. 267–297.

Hainmueller, J. and D. J. Hopkins (2014). “Public attitudes toward immigration”. In: *Annual Review of Political Science* 17, pp. 225–249.

Han, K. J. (2015). “The impact of radical right-wing parties on the positions of mainstream parties regarding multiculturalism”. In: *West European Politics* 38.3, pp. 557–576.

Hansen, K. M. and R. T. Pedersen (2015). “Negative campaigning in a multiparty system”. In: *Scandinavian Political Studies* 31.4, pp. 408–427.

Harmel, R. and L. Svåsand (1997). “The influence of new parties on old parties’ platforms: The cases of the Progress Parties and Conservative Parties of Denmark and Norway”. In: *Party Politics* 3.3, pp. 557–576.

Helbling, M. and A. Tresch (2011). “Measuring party positions and issue salience from media coverage: Discussing and cross-validating new indicators”. In: *Electoral Studies* 30.1, pp. 174–183.

Hellström, J. and M. Blomgren (2016). “Party debate over Europe in national election campaigns: Electoral disunity and party cohesion”. In: *European Journal of Political Research* 55.2, pp. 265–282.

Hetherington, M. J (2009). “Putting Polarization in Perspective”. In: *British Journal of Political Science* 39.2, pp. 413–448.

Hopkins, D. J. and G. King (2010). “A Method of Automated Nonparametric Content Analysis for Social Science”. In: *American Journal of Political Science* 54.1, pp. 229–247.

Huber, J. and R. Inglehart (1995). “Expert interpretations of party space and party locations in 42 societies”. In: *Party politics* 55.1, pp. 73–111.

Immerzeel, T., M. Lubbers, and H Coffé (2016). “Competing with the radical right: Distances between the European radical right and other parties on typical radical right issues”. In: *Party politics* 22.6, pp. 823–834.

— (2011). *Expert Judgement Survey of European Political Parties 2010*. URL: `https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset,%2055725`.

Inbar, Y. et al. (2012). “Disgust sensitivity, political conservatism, and voting”. In: *Social Psychological and Personality Science* 3.5, pp. 537–544.

Ivarsflaten, E. (2008). “What unites right-wing populists in Western Europe? Re-examining grievance mobilization models in seven successful cases”. In: *Comparative Political Studies* 41.1, pp. 3–23.

Jamieson, K. H. and J.N. Cappella (2008). *Echo Chamber: Rush Limbaugh and the Conservative Media Establishment*. New York: Oxford University Press.

Joulin, A. et al. (2016). “Fasttext.zip: Compressing Text Classification Models”. In: *arXiv preprint arXiv:1612.03651*.

Jungar, A. C. and A. R. Jupskås (2014). “Populist radical right parties in the Nordic region: A new and distinct party family?” In: *Scandinavian Political Studies* 37.3, pp. 215–238.

Kanaris, I., K. Kanaris, and E. Stamatatos (2006). “Spam Detection using Character N-grams”. In: *Advances in Artificial Intelligence*. Ed. by Antoniou G. et al. Heidelberg: SETN, pp. 95–104.

Kaplowitz, S. A. et al. (1983). “Anonymity, strength of attitude, and the influence of public opinion polls”. In: *Human Communication Research* 10.1, pp. 5–25.

Kernecker, T. and M. Wagner (2019). “Niche parties in Latin America”. In: *Journal of Elections, Public Opinion and Parties* 29.1, pp. 102–124.

Kitschelt, H. (2007). “Growth and persistence of the radical right in postindustrial democracies: Advances and challenges in comparative research”. In: *West European Politics* 30.5, pp. 1176–1206.

Kitschelt, H. and A. McGann (1995). *The radical right in Western Europe: A comparative analysis. The contemporary radical right: An interpretative and explanatory framework*, pp. 1–46.

König, T., M. Marbach, and M. Osnabrügge (2013). “Estimating party positions across countries and time—A dynamic latent variable model for manifesto data”. In: *Political analysis* 21.4, pp. 468–491.

Koopmans, R. and J. Muis (2009). “The rise of right-wing populist Pim Fortuyn in the Netherlands: A discursive opportunity approach”. In: *European Journal of Political Research* 48.5, pp. 642–664.

Kuran, T. (1989). “Sparks and prairie fires: A theory of unanticipated political revolution”. In: *Public choice* 61.1, pp. 41–74.

— (1991). “Now out of never: The element of surprise in the East European revolution of 1989”. In: *World politics* 44.1, pp. 7–48.

— (1995). “The inevitability of future revolutionary surprises”. In: *American Journal of Sociology* 100.6, pp. 1528–1551.

Lenz, G. S (2013). *Follow the Leader?: How Voters Respond to Politicians’ Policies and Performance*. Chicago: University of Chicago Press.

Lubbers, M., M. Gijsberts, and P. Scheepers (2002). “Extreme right-wing voting in Western Europe”. In: *European Journal of Political Research* 41.3, pp. 345–378.

Major, V., A. Surkis, and Y. Aphinyanaphongs (2017). "Utility of General and Specific Word Embeddings for Classifying Translational Stages of Research". In: *arXiv preprint arXiv:1705.06262*.

Mannes, J (2016). "Facebook's Artificial Intelligence Research Lab Releases Open Source fastText on GitHub". In: *Techcrunch*. URL: `https://techcrunch.com/2016/08/18/facebooks-artificial-intelligence-research-lab-releases-open-source-fasttext-on-github/`.

Mason, L (2015). "'I Disrespectfully Agree': The Differential Effects of Partisan Sorting on Social and Issue Polarization". In: *American Journal of Political Science* 59.1, pp. 128–145.

Mayer, N. (2015). "The closing of the radical right gender gap in France?" In: *French Politics* 13.4, pp. 391–414.

Meguid, B. M. (2005). "Competition between unequals: The role of mainstream party strategy in niche party success". In: *American Political Science Review* 99.3, pp. 347–359.

Minkenberg, M. (2000). "The renewal of the radical right: between modernity and anti-modernity". In: *Government and opposition* 35.2, pp. 170–188.

Mohammad, S. M. and P. D. Turney (2013). "Crowdsourcing a word–emotion association lexicon". In: *Computational Intelligence* 29.3, pp. 436–465.

Mudde, C. (2007). *Populist radical right parties in Europe*. United Kingdom: Cambridge University Press.

Munson, S. A., S. Y. Lee, and P. Resnick (2013). "Encouraging Reading of Diverse Political Viewpoints with a Browser Widget". In: *Seventh International AAAI Conference on Weblogs and Social Media (ICWSM)*.

Munson, S. A. and P. Resnick (2010). "Presenting Diverse Political Opinions: How and How Much". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1457–1466.

Mutz, D. C. (1989). "The influence of perceptions of media influence: Third person effects and the public expression of opinions". In: *International Journal of Public Opinion Research* 1.1, pp. 3–23.

— (1992). "Impersonal influence: Effects of representations of public opinion on political attitudes". In: *Political Behavior* 14.2, pp. 89–122.

Mutz, D. C. and P. S. Martin (2001). "Facilitating communication across lines of political difference: The role of mass media". In: *American Political Science Review* 95.1, pp. 97–114.

Newman, N. et al. (2017). "Reuters Institute Digital News Report 2017". In: *Reuters Institute for the Study of Journalism*.

Noelle-Neumann, E. (1974). "The spiral of silence a theory of public opinion". In: *Journal of communication* 24.2, pp. 43–51.

— (1989). "La spirale du silence: Une theorie de l'opinion publique [The spiral of silence: A theory of public opinion]". In: *Le Nouvel Espace Public: Hermes Cognition Communication Politique*. Paris: Editions du Centre National de la Recherche Scientifique, pp. 181–189.

Noelle-Neumann, E. (1995a). "Public opinion and rationality". In: *Public opinion and the communication of consent,* pp. 33–54.

— (1995b). *The spiral of silence: Public opinion, our social skin.* University of Chicago Press.

Norrander, B. (1991). "Explaining individual participation in presidential primaries". In: *Western Political Quarterly* 44.3, pp. 640–655.

Norris, P. (2005). *Radical right: Voters and parties in the electoral market.* United Kingdom: Cambridge University Press.

Nygård, M. and M. Kuisma (2017). "Immigration, integration and the Finns Party: Issue ownership by coincidence or by stealth?" In: *The European mainstream and the populist radical right.* Routledge, pp. 87–105.

Pariser, E (2011). *The Filter Bubble: What the Internet is Hiding from You.* United Kingdom: Penguin.

Park, S. et al. (2009). "NewsCube: Delivering Multiple Aspects of News to Mitigate Media Bias". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 443–452.

Patterson, T. E. (1980). *The mass media election: How Americans choose their president.* Praeger Publishers.

Pedersen, R. T. (2014). "News media framing of negative campaigning". In: *Mass Communication and Society* 17.6, pp. 898–919.

Perry, S. D. and W. J. Gonzenbach (1997). "Effects of news exemplification extended: Considerations of controversiality and perceived future opinion". In: *Journal of Broadcasting and Electronic Media* 41.2, pp. 229–244.

— (2000). "Inhibiting speech through exemplar distribution: Can we predict a spiral of silence?" In: *Journal of Broadcasting and Electronic Media* 44.2, pp. 268–281.

Petty, R. E. and J. T. Cacioppo (1979). "Issue involvement can increase or decrease persuasion by enhancing message-relevant cognitive responses". In: *Journal of personality and social psychology* 37.10, pp. 1915–1926.

— (1981). "Personal involvement as a determinant of argument-based persuasion". In: *Journal of personality and social psychology* 41.5, pp. 847–855.

Poole, K. T. and H. Rosenthal (1984). "The Polarization of American Politics". In: *The Journal of Politics* 46.4, pp. 1061–1079.

Prior, M (2013). "Media and Political Polarization". In: *Annual Review of Political Science* 16, pp. 101–127.

Redlawsk, D. P., A. Civettini, and K. M. Emmerson (2010). "The affective tipping point: Do motivated reasoners ever get it?" In: *Political Psychology* 31, pp. 563–593.

Roberts, M. E., B. M. Stewart, and D. Tingley (2014). "stm: R package for structural topic models". In: *Journal of Statistical Software* 10.2, pp. 1–40.

Rovny, J. (2013). "Where do radical right parties stand? Position blurring in multidimensional competition". In: *European Political Science Review* 5.1, pp. 1–26.

Ruedin, D. (2013). "Obtaining party positions on immigration in Switzerland: Comparing different methods". In: *Swiss Political Science Review* 19.1, pp. 84–105.

Rydgren, J. (2008). "Immigration sceptics, xenophobes or racists? Radical right-wing voting in six West European countries". In: *European Journal of Political Research* 47.6, pp. 737–765.

— (2012). "Introduction: Class politics and the radical right." In: *Class politics and the radical right*. Routledge, pp. 19–27.

Scheufele, D. A. and P. Moy (2000). "Twenty-five years of the spiral of silence: A conceptual review and empirical outlook". In: *International journal of public opinion research* 12.1, pp. 3–28.

Scheufele, D. A., J. Shanahan, and E Lee (2001). "Real talk: Manipulating the dependent variable in spiral of silence research". In: *Communication research* 28.3, pp. 304–324.

Shields, J. (2007). *The extreme right in France: from Pétain to Le Pen*. Routledge.

Silva, C. C. et al. (2018). "Extracting In-domain Training Corpora for Neural Machine Translation Using Data Selection Methods". In: *Proceedings of the Third Conference on Machine Translation: Research Papers*, pp. 224–231.

Slapin, J. B. and S. O. Proksch (2008). "A scaling model for estimating time-series party positions from texts". In: *American Journal of Political Science* 52.3, pp. 705–722.

Smith, J. M. (2010). "Does crime pay? Issue ownership, political opportunity, and the populist right in Western Europe". In: *Comparative Political Studies* 43.11, pp. 1471–1498.

Sonnichsen, N. (2019). *Total number of Facebook users in Germany in January 2017, by age group and gender (in millions)*. URL: https://www.statista.com/statistics/867622/number-of-facebook-users-in-germany-by-age-and-gender/.

Stiftung, B. (Ed.) (2018). *Transformation Index BTI 2018: Governance in International Comparison*.

Stroud, N. J. (2008). "Media use and political predispositions: Revisiting the concept of selective exposure". In: *Political Behavior* 30.3, pp. 341–366.

Sunstein, C. R. (2009). *Going to Extremes: How Like Minds Unite and Divide*. Oxford: Oxford University Press.

— (2018). *Republic: Divided Democracy in the Age of Social Media*. Princeton: Princeton University Press.

Swank, D. and H. G. Betz (2003). "Globalization, the welfare state and right-wing populism in Western Europe". In: *Socio-Economic Review* 1.2, pp. 215–245.

Szöcsik, E. and A. Polyakova (2019). "Euroscepticism and the electoral success of the far right: the role of the strategic interaction between center and far right". In: *European political science* 18.3, pp. 400–420.

Tartar, A. (2017). *How the Populist Right Is Redrawing the Map of Europe*. URL: https://www.bloomberg.com/graphics/2017-europe-populist-right/.

Van der Brug, W. (2004). "Issue ownership and party choice". In: *Electoral studies* 23.2, pp. 209–233.

Van Spanje, J. (2010). "Contagious parties: Anti-immigration parties and their impact on other parties' immigration stances in contemporary Western Europe". In: *Party Politics* 16.5, pp. 563–586.

Van Spanje, J. and C. de Vreese (2014). “Europhile media and Eurosceptic voting: Effects of news media coverage on Eurosceptic voting in the 2009 European parliamentary elections”. In: *Political Communication* 31.2, pp. 325–354.

Vasilopoulos, P. (2018). “Terrorist events, emotional reactions, and political participation: The 2015 Paris attacks”. In: *European Politics* 41.1, pp. 102–127.

Vasilopoulos, P. et al. (2019). “Fear, anger, and voting for the far right: Evidence from the November 13, 2015 Paris terror attacks”. In: *Political Psychology* 40.4, pp. 679–704.

Vlandas, T. and D. Halikiopoulou (2019). “Does unemployment matter? Economic insecurity, labour market policies and the far-right vote in Europe”. In: *European Political Science* 18.3, pp. 421–438.

Yardi, S. and D. Boyd (2010). “Dynamic Debates: An Analysis of Group Polarization Over Time on Twitter”. In: *Bulletin of Science, Technology and Society* 30.5, pp. 316–327.

Yin, C. C. (1998). “Equilibria of collective action in different distributions of protest thresholds”. In: *Public Choice* 97.4, pp. 535–567.

Young, L. and S. Soroka (2012). “Affective news: The automated coding of sentiment in political texts”. In: *Political Communication* 29.2, pp. 205–231.

Zhang, L. and B. Liu (2011). “Identifying Noun Product Features that Imply Opinions”. In: *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers* 2, pp. 575–580.

Appendix A

Contingent Extremism: How Perceptions of Party Popularity Activate Far Right Support

A.1 Supplemental figures and tables, including information on the list experiment

	Hungary		France		Germany	
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
High polling treatment	1.16*	1.18*	1.23**	1.26**	1.36***	1.37***
	(0.10)	(0.11)	(0.11)	(0.11)	(0.11)	(0.12)
Favorable opinion climate		1.35**		1.50***		2.98***
		(0.18)		(0.19)		(0.62)
Male		1.05		1.05		1.64***
		(0.14)		(0.14)		(0.20)
HS diploma		0.99		0.93		1.05
		(0.17)		(0.19)		(0.26)
Some college		0.89		0.75		0.68**
		(0.21)		(0.15)		(0.13)
College degree		0.62**		0.53*		0.61***
		(0.14)		(0.18)		(0.11)
Graduate degree		0.46***		0.54**		1.38*
		(0.12)		(0.14)		(0.25)
Age 25-34		1.30		0.88		1.49**
		(0.39)		(0.25)		(0.29)
Age 35-44		1.04		1.40		2.38***
		(0.28)		(0.30)		(0.53)
Age 45-54		0.83		0.98		2.48***
		(0.22)		(0.19)		(0.47)
Age 55-64		0.64*		1.08		2.76***
		(0.16)		(0.21)		(0.54)
Age 65+		0.83		1.10		3.40***
		(0.24)		(0.26)		(0.76)
Catholic		1.29		0.92		1.44*
		(0.43)		(0.34)		(0.30)
Jewish		0.83		0.00		1.82
		(1.08)		(0.00)		(0.95)
Muslim		0.93		0.33*		0.60
		(0.92)		(0.20)		(0.41)
Non-religious		0.96		0.29***		1.04
		(0.33)		(0.12)		(0.18)
Other religion		1.34		0.50*		1.32
		(0.50)		(0.18)		(0.33)
Num. obs.	1015	1015	1770	1770	1991	1991

Table A.1: Supplementary Robustness Checks on the Predicted Probability of Far Right Identification (Figure 2.5)

This table supplements Figure 2.5 in the main text. Models 1, 3, and 5 are the logit models that form the basis for the predicted probabilities. This table reports odds ratios for all logit models. Models 1, 3, and 5 use a logit specification to predict whether an individual will identify as a far right supporter (1) or not (0), and includes our main independent variable of interest (the polling treatment). Models 2, 4, and 6 also include controls for: favorability of regional opinion climate, gender, level of education, age, and religion.

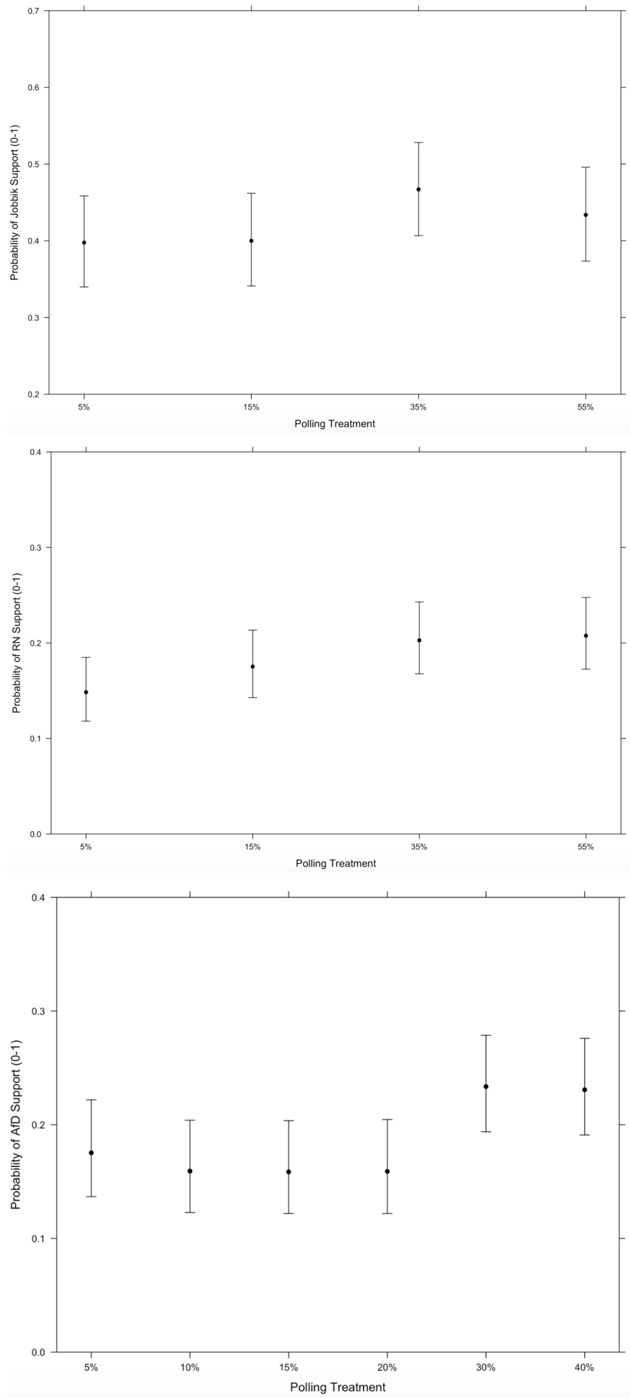


Figure A.1: Predicted Probability of Far Right Identification by Unpooled Polling Treatment (with 95% CIs) (Figure 2.5)

These figures serve as a supplement to Figure 2.5 in the main text. Logistic regression predicting far right identification, by unpooled polling treatment. From top to bottom: Hungary, France, Germany.

	List Experiment (Raw)	List Experiment (Imputed)	Low Polls	High Polls
Hungary				
Control	0.99 [367]	0.99 [367]		
Treatment	1.41 [336]	1.458 [367]		
% Far Right	42%	46.8%	39.9%	45.1%
N	703	734	509	506
France				
Control	1.27 [611]	1.27 [611]		
Treatment	1.37 [549]	1.461 [611]		
% Far Right	10%	19.1%	16.2%	20.5%
N	1160	1222	883	887
Germany				
Control	1.15 [805]	1.15 [805]		
Treatment	1.36 [796]	1.369 [805]		
% Far Right	21%	21.9%	16.3%	23.2%
N	1601	1610	1233	758

Table A.2: Far Right Identification, Direct and Unobtrusive Measures (Explaining the Failed List Experiment)

Far right identification rates (means and 95% CIs) in polling treatments (direct measure) and in the list experiment (unobtrusive measure). The variance of the list experiment estimator is calculated with the standard large-sample formula for a difference-in-means. As Table A.2 illustrates, the list experiment leads to estimates with high variance when compared to direct questions because the sensitive item (far right identification) is aggregated with non-sensitive items. Although high variance is the cost of protecting respondents' anonymity, this feature makes the list experiment unreliable as a 'check' on the ceiling of far right support. Moreover, even though the list experiment is designed to provide privacy, survey respondents appear to have been highly sensitive to answering it truthfully. Although the list experiment was properly randomized, a significantly higher proportion of respondents refused to answer the list experimental question in the treatment condition (featuring the far right party) than in the control (excluding the far right party). This compliance gap is consistent across all three countries, and may explain why the list experimental estimates of far right support are unexpectedly low–especially in France. Moreover, chronologically, the list experiment proceeded the polling treatment in the survey. This chronology may have led respondents to 'figure out' the objective of the list experiment, and thereby feel more sensitive to identifying as a far right supporter than under typical conditions of anonymity.

	Hungary		France		Germany	
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
High polling treatment	1.27**	1.31**	1.34**	1.36**	1.38***	1.37***
	(0.15)	(0.16)	(0.16)	(0.16)	(0.12)	(0.12)
Favorable opinion climate	1.39**	1.35**	1.56***	1.52***	2.39***	2.96***
	(0.18)	(0.18)	(0.19)	(0.19)	(0.48)	(0.63)
Polling treatment*Favorable	0.80	0.77	0.84	0.84	0.90	0.96
	(0.15)	(0.15)	(0.15)	(0.15)	(0.25)	(0.28)
Male		1.05		1.04		1.64***
		(0.14)		(0.14)		(0.20)
HS diploma		0.99		0.94		1.05
		(0.17)		(0.19)		(0.26)
Some college		0.90		0.75		0.68**
		(0.22)		(0.15)		(0.13)
College degree		0.62**		0.55*		0.61***
		(0.14)		(0.18)		(0.11)
Graduate degree		0.46***		0.54**		1.38*
		(0.12)		(0.14)		(0.25)
Age 25-34		1.31		0.87		1.49**
		(0.40)		(0.25)		(0.29)
Age 35-44		1.03		1.39		2.39***
		(0.28)		(0.30)		(0.53)
Age 45-54		0.83		0.98		2.48***
		(0.22)		(0.19)		(0.47)
Age 55-64		0.64*		1.08		2.76***
		(0.16)		(0.21)		(0.54)
Age 65+		0.83		1.10		3.40***
		(0.24)		(0.26)		(0.76)
Catholic		1.31		0.91		1.44*
		(0.44)		(0.34)		(0.30)
Jewish		0.81		0.00		1.83
		(1.05)		(0.00)		(0.96)
Muslim		0.93		0.33*		0.61
		(0.92)		(0.20)		(0.41)
Non-religious		0.97		0.29***		1.04
		(0.33)		(0.11)		(0.18)
Other religion		1.39		0.49*		1.32
		(0.52)		(0.18)		(0.33)
Num. obs.	1015	1015	1770	1770	1991	1991

Table A.3: Supplementary Robustness Checks on the Predicted Probability of Far Right Identification by Favorability of Opinion Climate (Figure 2.6)

This table supplements Figure 2.6, reporting odds ratios for all logit models. Models 1, 3, and 5 use a logit specification to predict whether an individual will identify as a far right supporter (1) or not (0), depending on whether they reside in a regionally favorable opinion climate (1) or not (0). Models 2, 4, and 6 also include controls for: gender, level of education, age, and religion.

Appendix B

Examining Far Right Campaign Appeals through Digital Political Advertising

B.1 Sample stm diagnostics by number of topics

B.2 Supplementary topic prevalence contrast plots (Table 3.2)

These country plots illustrate the change in topic proportions shifting from far right party documents to other party documents. The top 'FREX' words are displayed, meaning words which are both frequent in and exclusive to a topic of interest. As the Table 3.2 summary of this content shows, most far right parties have clear issue ownership over immigration-related topics.

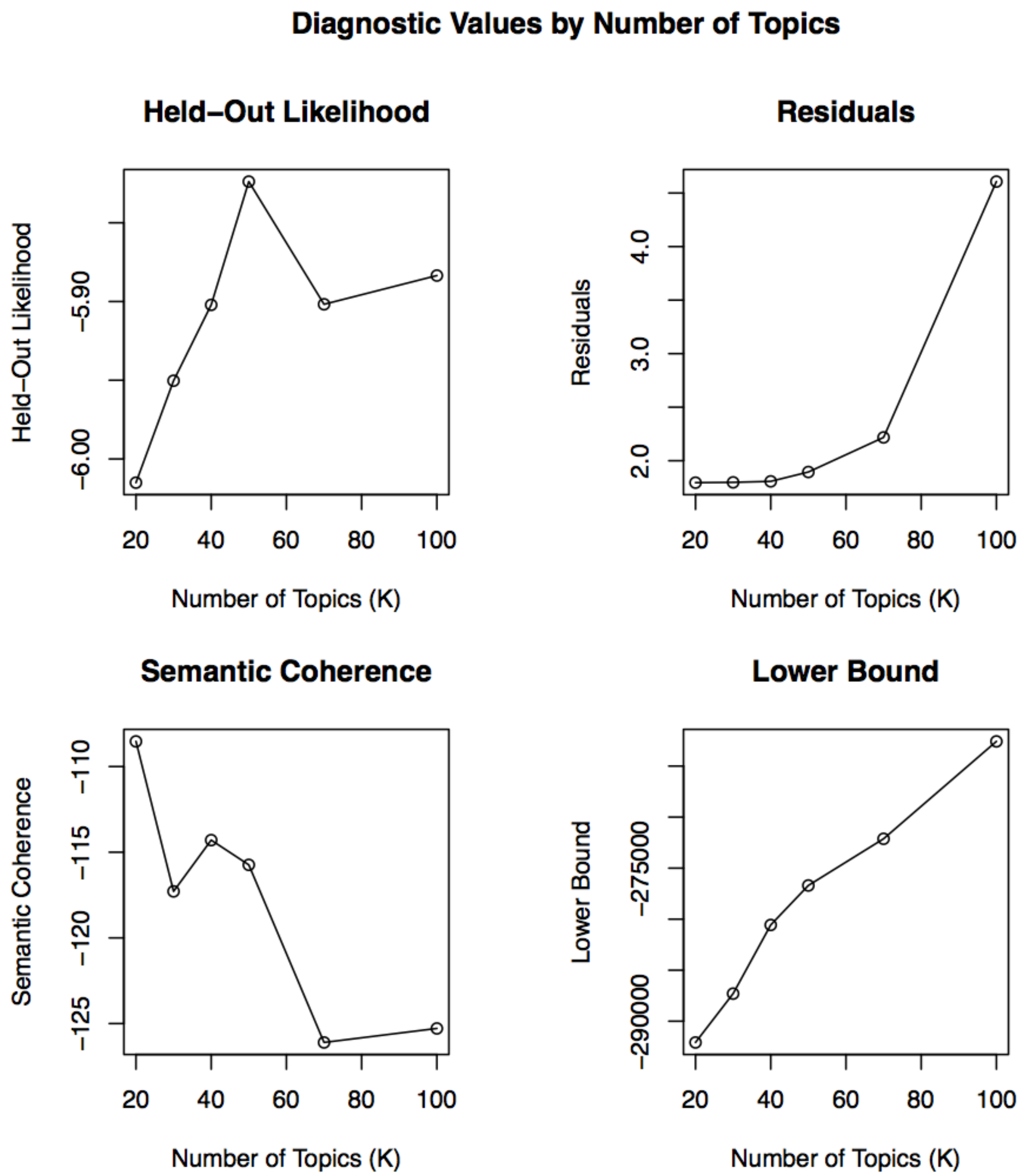


Figure B.1: Sample Stm Diagnostics by Number of Topics

*

To choose an appropriate number of topics for the stm topic model, it is important to first evaluate model diagnostics on the residuals, the semantic coherence of the topics, and the held-out likelihood. In this example (from the German ad text), k=40 does a good job maximizing a combination of high semantic coherence, high held-out likelihood, and low residuals.

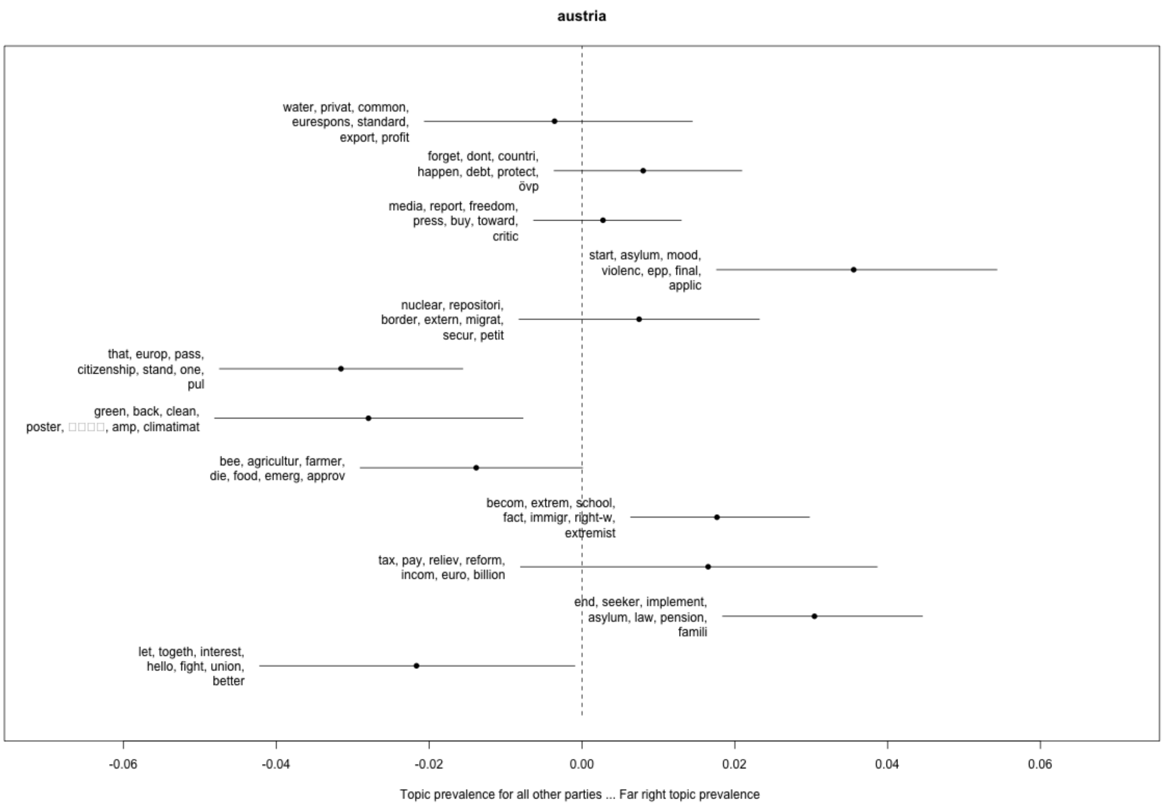


Figure B.2: Topical Prevalence Contrast (Austria)

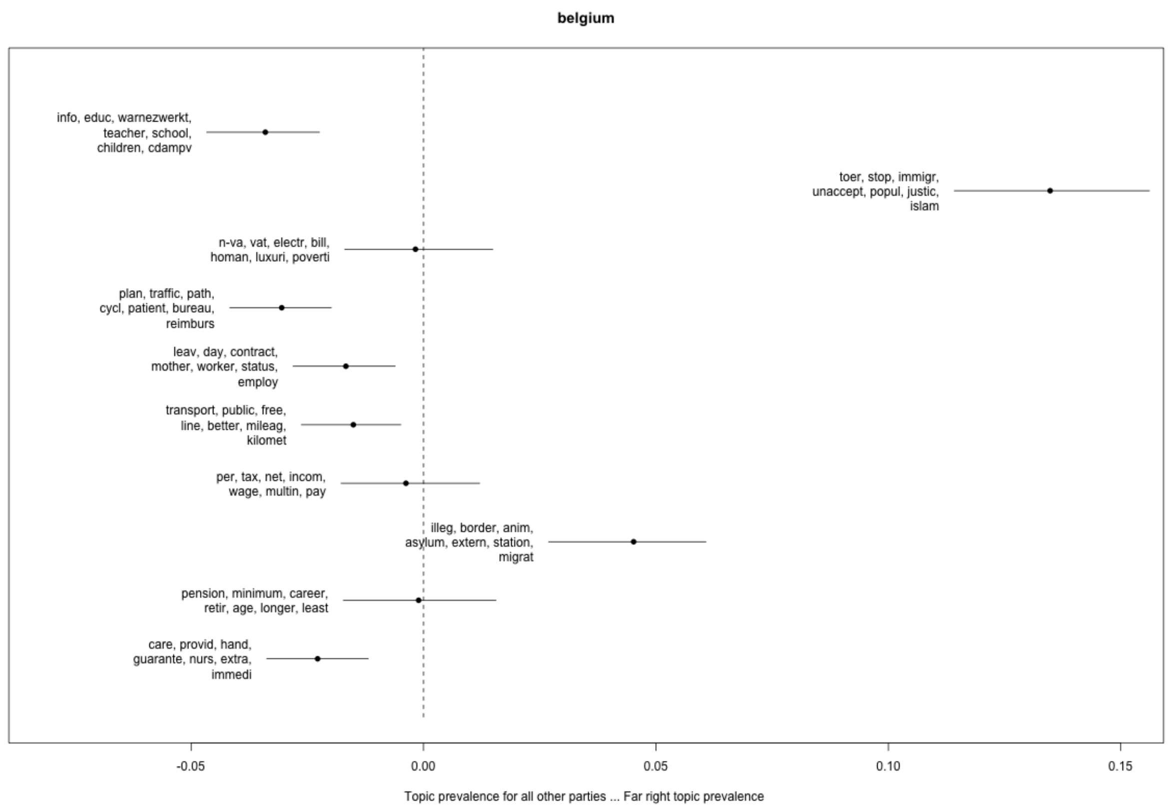


Figure B.3: Topical Prevalence Contrast (Belgium)

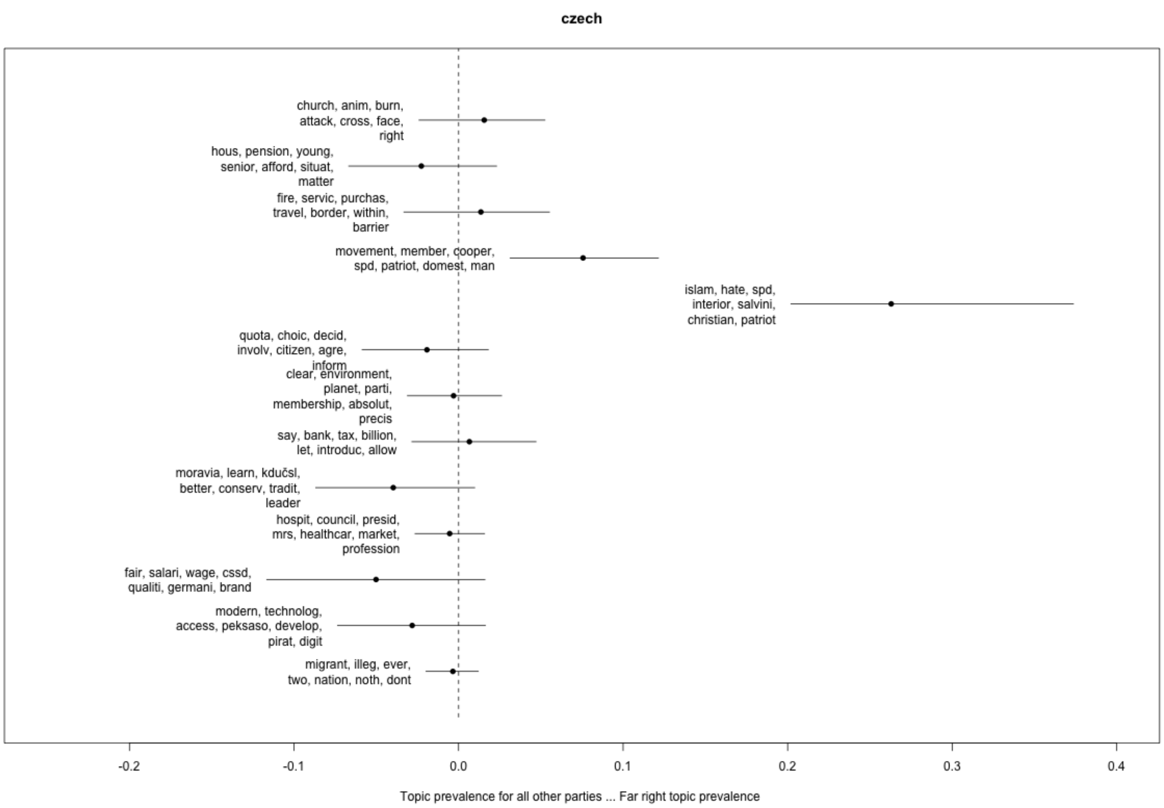


Figure B.4: Topical Prevalence Contrast (Czech Republic)

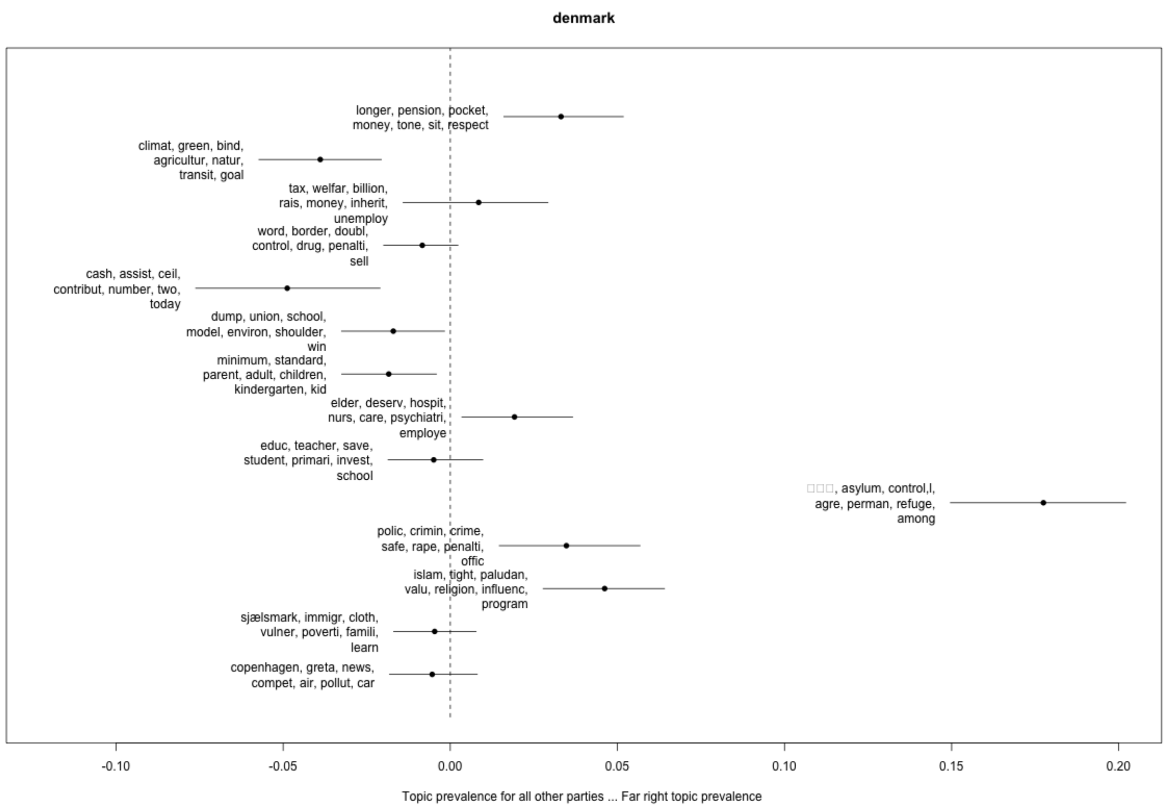


Figure B.5: Topical Prevalence Contrast (Denmark)

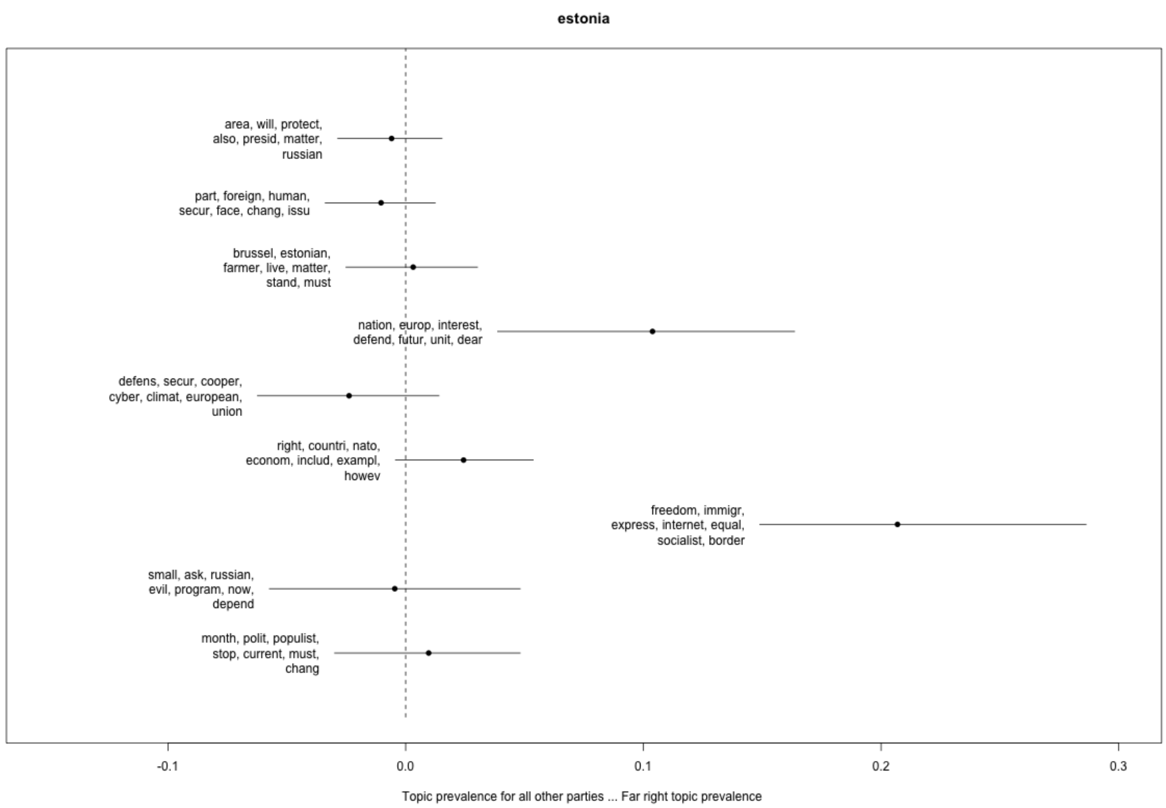


Figure B.6: Topical Prevalence Contrast (Estonia)

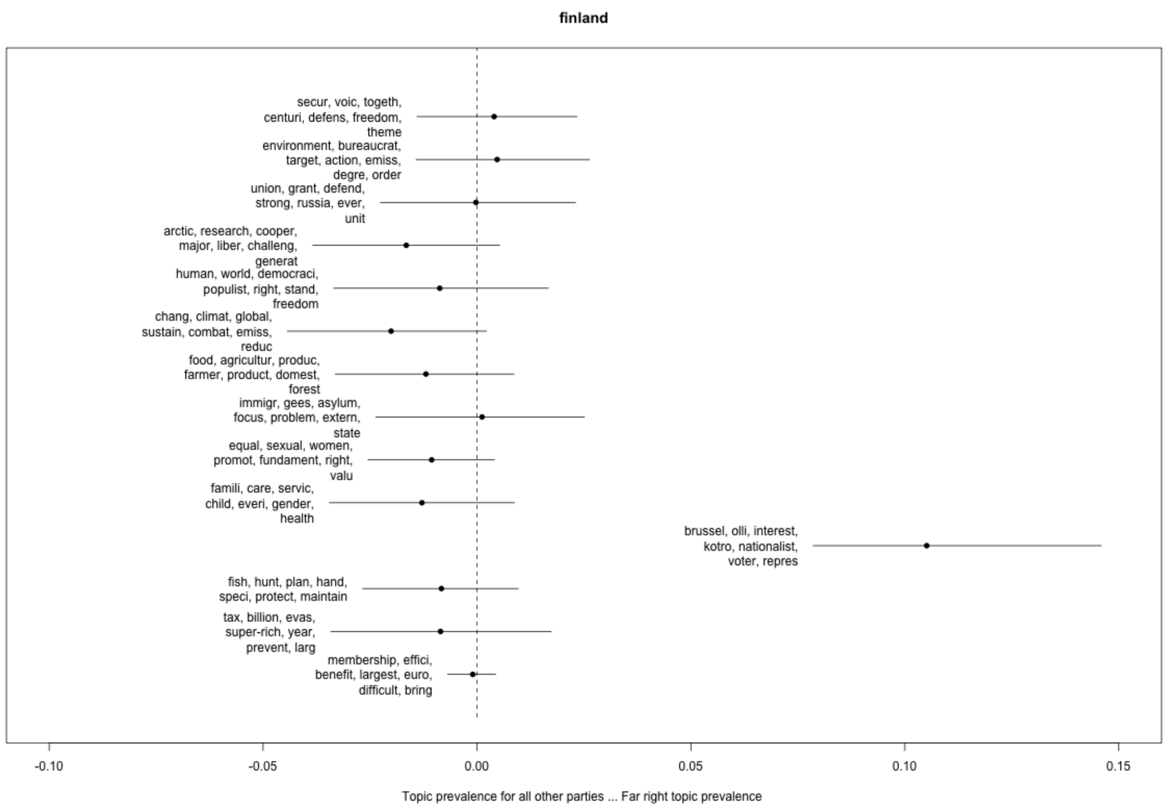
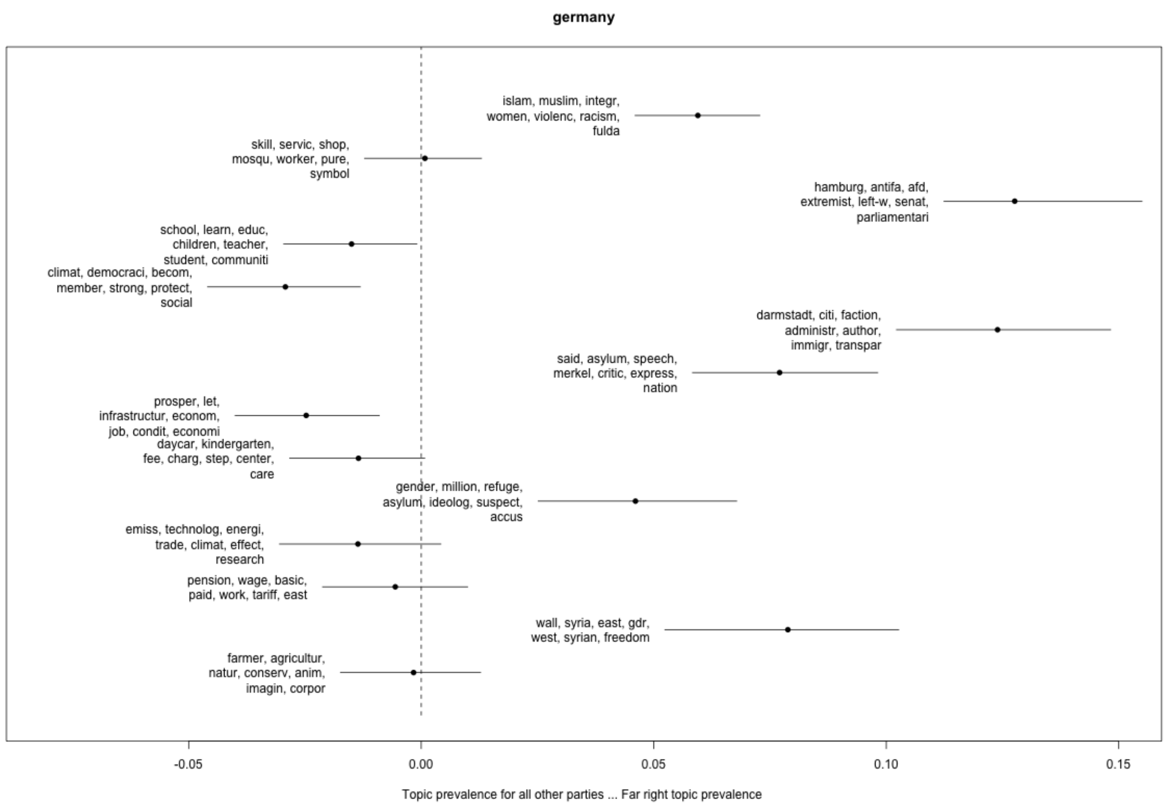


Figure B.7: Topical Prevalence Contrast (Finland)



Topical Prevalence Contrast (Germany)

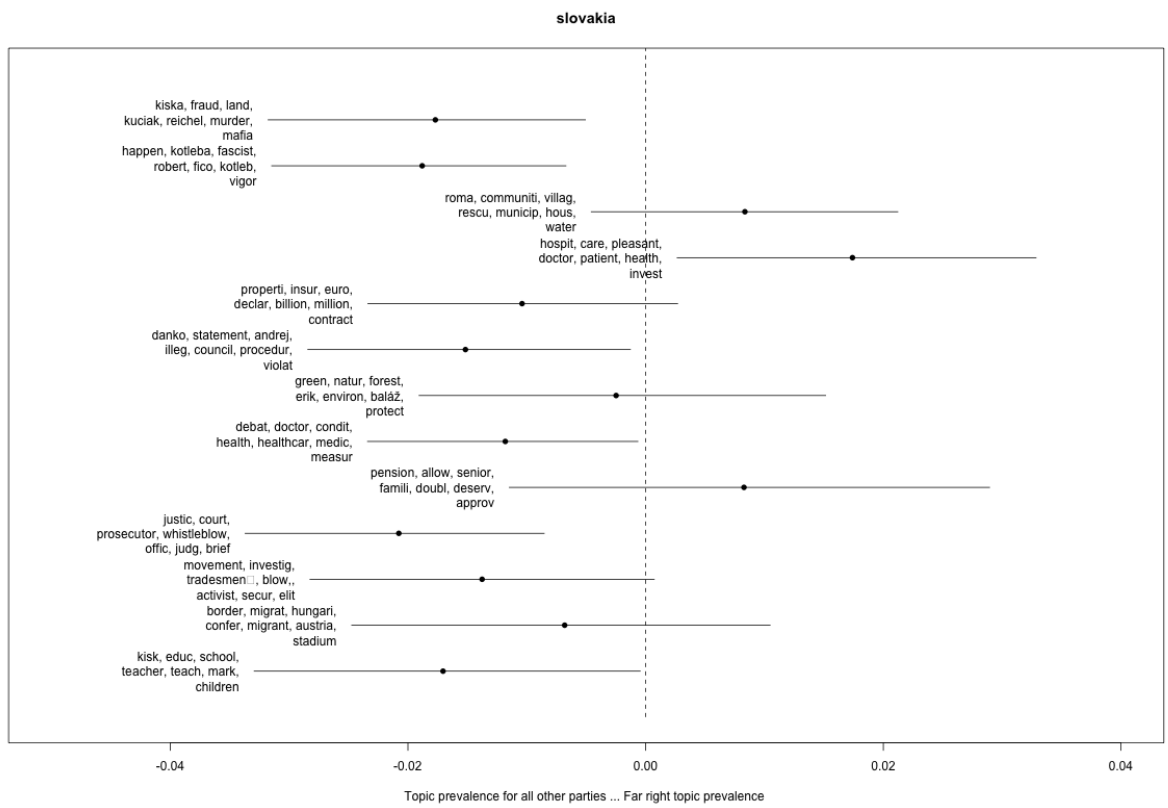


Figure B.8: Topical Prevalence Contrast (Slovakia)

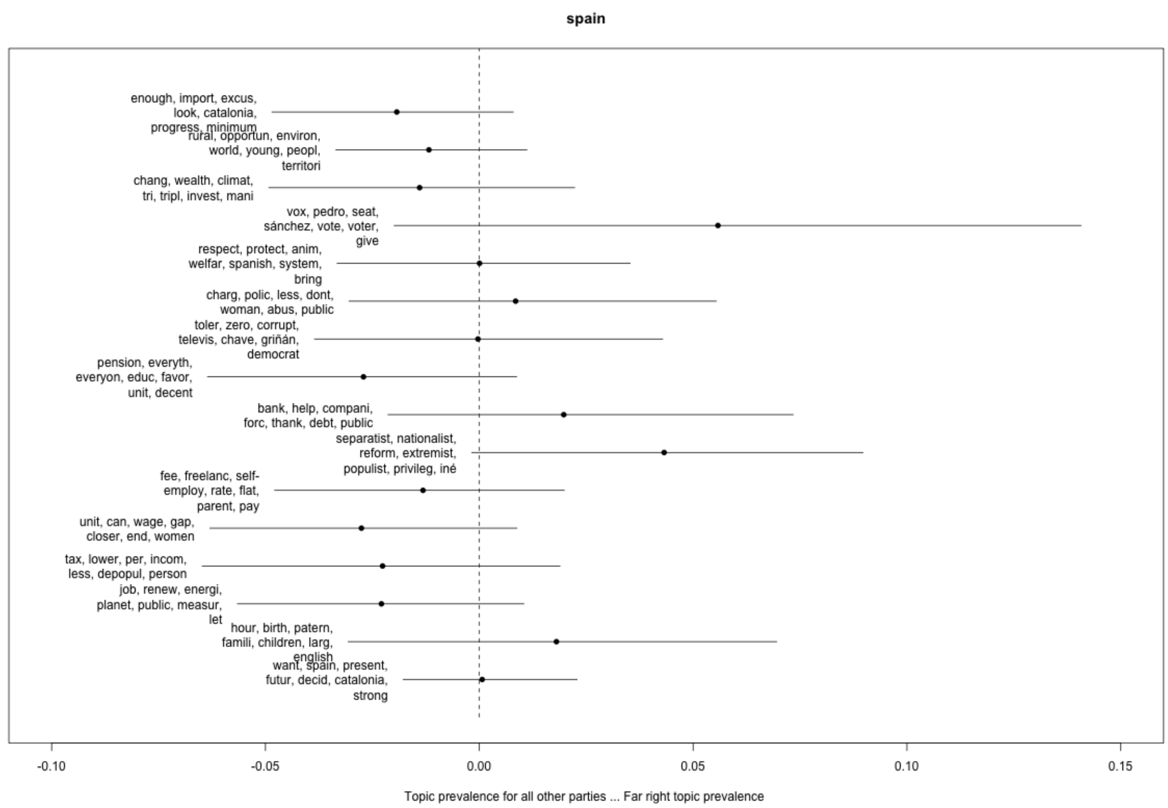


Figure B.9: Topical Prevalence Contrast (Spain)

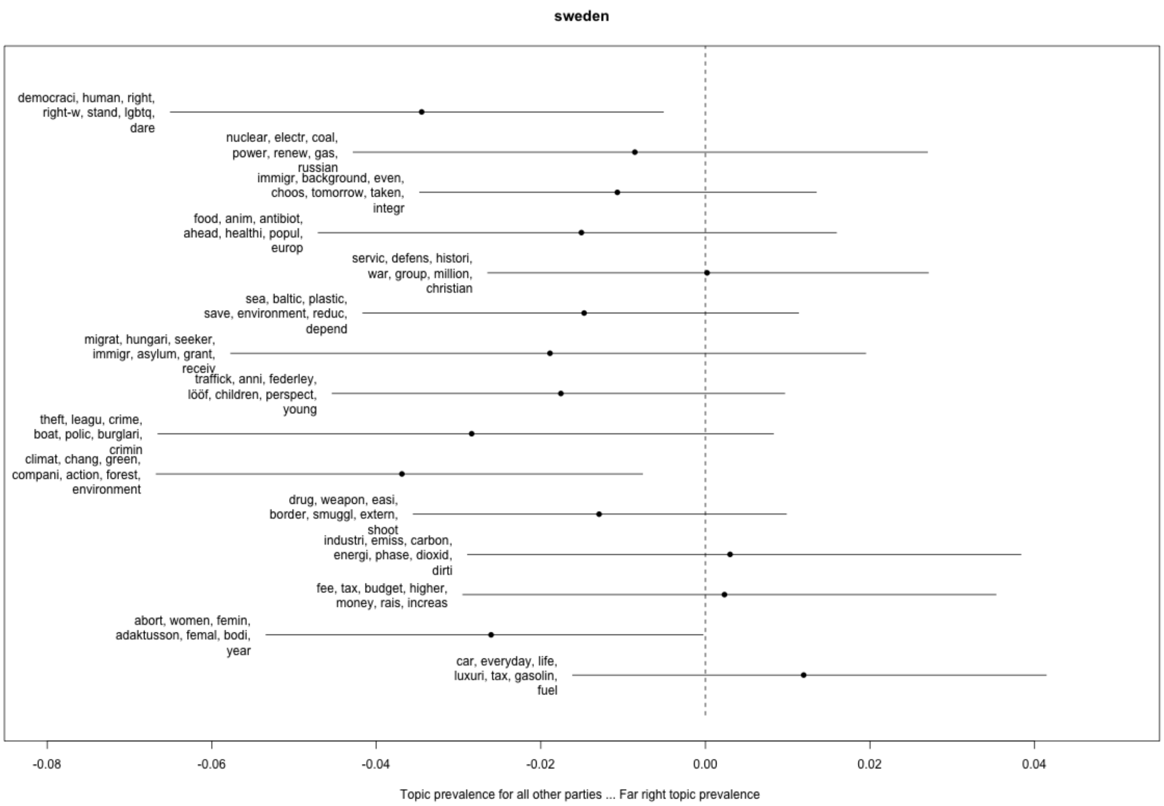


Figure B.10: Topical Prevalence Contrast (Sweden)

Appendix C

A Machine Learning Approach to Detecting Polarization in Social Media Discourse

C.1 Supplementary information on the research procedure

A. Data Collection

To create our training data, we used the standard API to sample content from media elites' user timelines. For our unlabelled data, we used the full-archive Search Tweets API to sample thousands of tweets on different political topics chronologically preceding and following the timing of highly politicized news events.

Twitter Search Parameters (by Event)

Note that some time frames to capture tweets are larger than others by a few days. This is because we needed to adjust the keyword search window until we could sample 5,000 tweets, and usually post-event, we hit the 5,000 tweet threshold in a shorter time frame.

Event No. 1: 2017 Las Vegas Mass Shooting
Pre-event sample time parameters: 2017-09-21 to 2017-09-30
3000 tweets featuring the keyword "gun control"; 358 tweets featuring the keyword "gun debate"; 2134 tweets featuring the keyword "gun rights"

Post-event sample time parameters: 2017-10-01 to 2017-10-04
3000 tweets featuring the keyword "gun control"; 360 tweets featuring the keyword "gun debate"; 2100 tweets featuring the keyword "gun rights"

Event No. 2: 2016 Brexit Referendum
Pre-event sample time parameters: 2016-06-21 to 2016-06-22
5000 tweets featuring the keyword "Brexit"

Post-event sample time parameters: 2016-06-24 to 2016-06-25
5000 tweets featuring the keyword "Brexit"

Event No. 3: 2016 U.S. Presidential Election
Pre-event sample time parameters: 2016-11-05 to 2016-11-07
2500 tweets featuring the keyword "Clinton"; 2500 tweets featuring the keyword "Trump"

Post-event sample time parameters: 2016-11-09 to 2016-11-11
2500 tweets featuring the keyword "Clinton"; 2500 tweets featuring the keyword "Trump"

B. Data Processing

Tweet Cleaning Procedure

• Replace shortened URLs with extended URLs (provided by the Twitter API)
(https://developer.twitter.com/en/docs/tweets/data-dictionary/overview/tweet-object)
• Remove duplicate tweets
• Remove control characters
• Replace new line and carriage returns with spaces
• Make all letters lower case
• Remove links to other twitter statuses
(links that start with https://twitter.com/i/web/status or https://twitter.com)
• For links starting with "https://" and "http://", replace punctuation with spaces
• For words with punctuation, remove punctuation (e.g., can't → cant)
• Remove tweets shorter than 5 words

fastText Parameters

For a full description and default fastText parameters, please reference:
https://fasttext.cc/docs/en/options.html

To improve the performance of the algorithm, there are two parameters we set that deviate from the defaults listed at the link. These are:
• wordNgrams=2
• learning_rate = 0.3

Additional Training Parameters

To create our training set, we select an equally sized tweet corpus for each label category. The sample size is the number of tweets from the smallest label category. For example, if the elites we sampled for 'far left' comprised just 60,000 tweets whereas the elites for the 'moderate' and 'far right' classification comprised 70,000 tweets, we would sample 60,000 tweets from each category (far left, far right, moderate) to create our training set for each bootstrap run.

Once we have this equalized training set, we split the data into 'training' and 'test' categories so that we can evaluate the performance of our algorithm.

- Percent of data used for training = 80
- Percent of data used for testing = 20

C. Elite's Ideology Scores, by Text Classification

We used Barbera's Bayesian spatial model to sample media elites' political media networks and proxy for their ideology. This offered a way to 'check' the validity of the text labels we imposed on our training data, and helped us adjust our elite labels accordingly.

Three observations are important to note. First, there are a few elites in each event sample that appear extreme along the ideological spectrum but receive a moderate classification label, and vice versa. These are elites that clearly represent the ideological camp we have labelled them, but who have embedded themselves in unexpected networks—either to engage in discourse with their political opposition, or to gain legitimacy with mainstream media, or for other strategic purposes. After reviewing the contents of their tweets manually, we believe we have assigned them an appropriate text classification.

Second, there are fewer elites sampled for Brexit than for the Vegas shooting and Trump's election. This is because a smaller core group of media elites discussed the topic of Brexit extensively enough to be representative of the relevant Twitter discussion. Moreover, Twitter deleted numerous Brexit-related media accounts over the duration of our study, due to suspected coordinated inauthentic behavior and foreign influence linked to those accounts.

Third, there are a number of far left and far right media elites that are not listed on the y-axis in Figures C.1–C.3 because they do not follow any traditional media accounts themselves and thus couldn't be scored. We do not exclude them from the training data, because it is relatively common for extreme media elites to only embed in alt-networks that would preclude them from being scored by Barbera's tweetscores package.

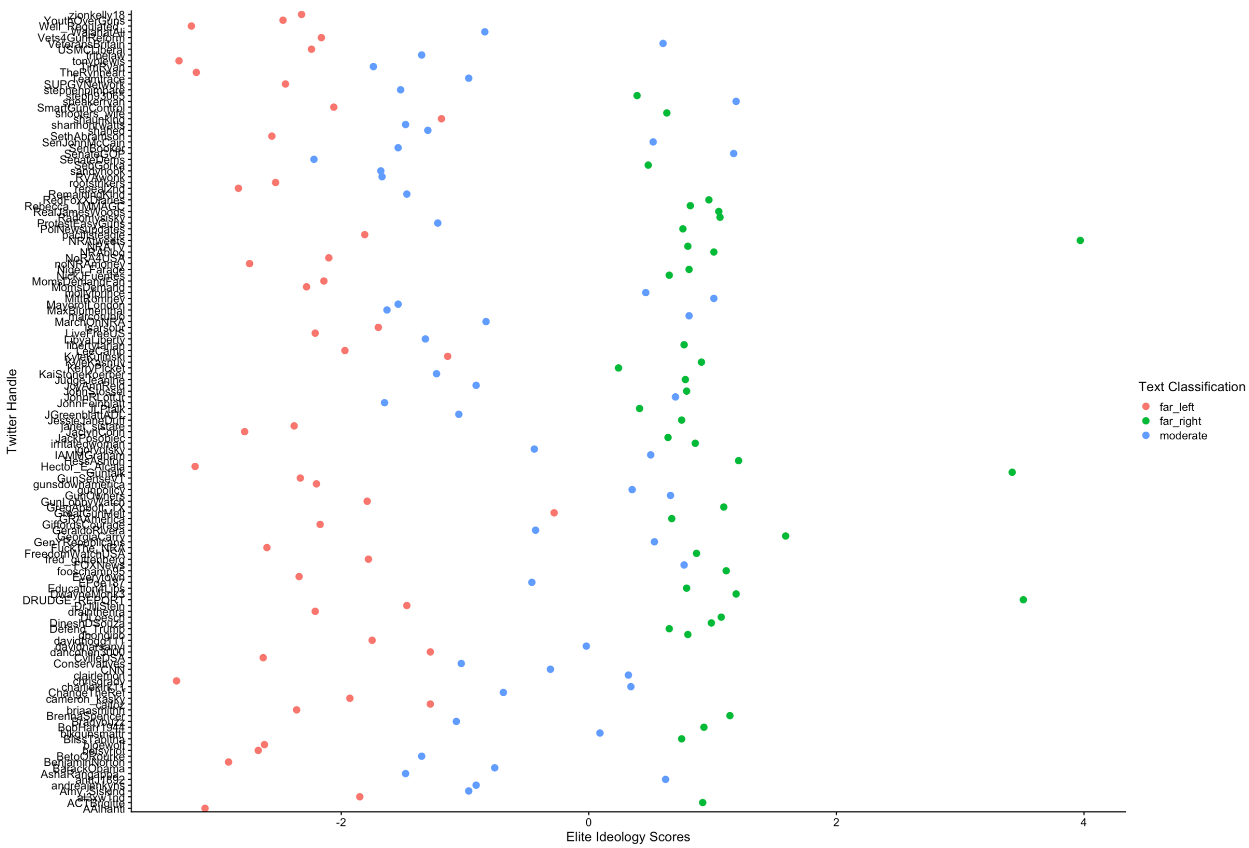


Figure C.1: Elite Ideology Scores for the 2017 Las Vegas Mass Shooting Training Data

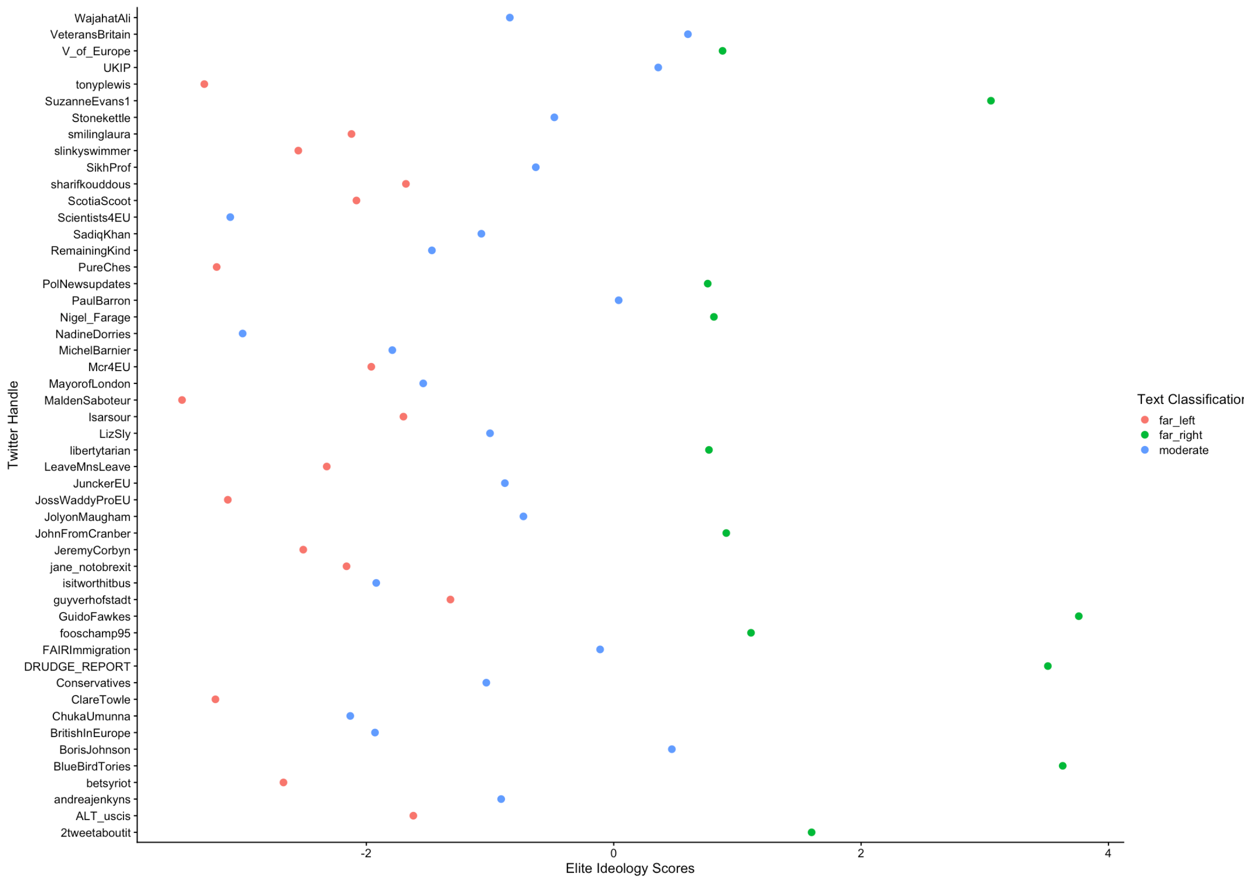


Figure C.2: Elite Ideology Scores for the 2016 Brexit Referendum Training Data

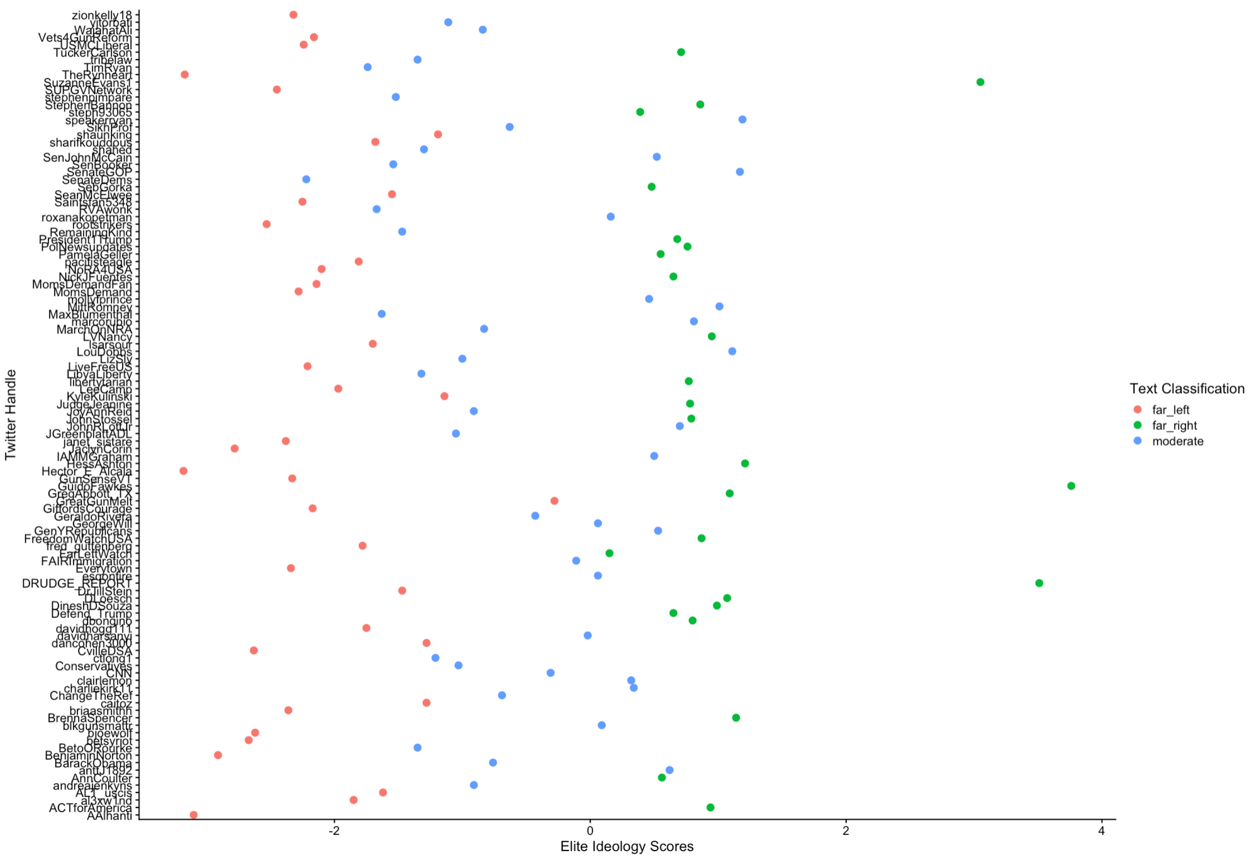


Figure C.3: Elite Ideology Scores for the 2016 U.S. Presidential Election Training Data