Terrorist Attacks, Cultural Incidents and the Vote for Radical Parties

Giavazzi, F.*, Iglhaut, F.*, Lemoli, G.† and Rubera, G.*

February 25, 2019

Abstract

We study the recent increase in support for a radical right-wing party in Germany (AfD) relating it to the cultural backlash associated with a number of terrorist attacks, which happened in Europe around the same time, and a very salient incident occurred in the city of Cologne on New Year's Eve 2015/2016. We measure shifts in voters' attitudes at the time of those events by the change in a measure of linguistic distance between the content of the tweets they posted and those posted by the main German parties. Next we use the date of those events to identify exogenous shifts in language similarity at the electoral constituency level, using a random coefficients model. We find that following such events the similarity with the AfD tweets increases while it increases much less or even falls for the other parties. We then regress differences in vote shares between the 2013 and 2017 general elections at the electoral constituency level on exogenous shifts in language similarity, finding that these are significant in explaining changes in vote shares.

*Bocconi University
†New York University

1 Introduction

In recent years in many Western democracies we have observed a shift away from the traditional principles of an open society. The very ideals on which such democracies were founded, including multi-culturalism, have been called into question. Alongside, the salience of immigration and national security in public discourse has increased. These dynamics have moved in parallel with the support for radical right-wing parties, and the corresponding decline of long-established moderate ones.

Common explanations for these phenomena point to economic shocks (the Great Recession) and to globalization, immigration flows and the increase in inequality, all of which have increased the malaise of specific groups in society. Globalization and immigration in turn increase the salience of cultural differences and raise the perception of their threats. The perceived distance between cultural norms raises concerns about the quantity of immigrants and increases the demand for stronger borders. Parties which emphasize such threats and promise policies to address them are bound to gain.

We investigate these phenomena studying one specific country and period: Germany between the last two federal elections, that is between 2013 and 2017. In this four year period the German economy was recovering from the Great Recession faster than other Eurozone countries: growth averaged 2 percent between 2014 and 2017 and unemployment fell from 7 percent to below 6 percent although with important differences across regions. Thus in Germany the effects of the economic crisis should have been considerably less salient than in other European countries. Importantly, however, the concerns about immigrants mentioned above were highlighted by two sets of events. First, a number of terrorist attacks, claimed mainly by the Islamic State of Iraq and the Levant (ISIL), hit various parts of Europe in those years: in Germany there were a total of seven attacks[1], the one on the Berlin Christmas market on December 15, 2016 with a very high death toll: 12 people died and 56 were seriously injured. Other non-terrorist incidents, happened in the same period: the most salient was the sexual assault on a number of German women on New Year's eve 2016. In the Cologne city center hundreds of women were victims of sex offences recorded by the police and similar incidents took place in other German cities: Hamburg, Dortmund, Düsseldorf, and Stuttgart. For all of Germany, police estimated that 1,200 women were sexually assaulted that night and at least 2,000 men were involved.[2] Cologne police chief Wolfgang Albers stated that the perpetrators in his city were reportedly men of "Arab or North African appearance" and said that Germany had never experienced such mass sexual assaults before.[3]

The second event is the migration of refugees from the Middle East, especially out of Syria and its civil war. In September 2015, Austria and Germany took one of the most pivotal decisions of Europe's refugee crisis, opening their borders to tens of thousands of migrants who had piled up in Hungary. A turning point in the political discourse was a speech by Chancellor Angela Merkel on August 31st 2015 when she said "I put it simply, Germany is a strong country ...we have managed so many things, we can do this", which was interpreted as an open door to refugees.

At the end of this 4-year period in the 2017 general elections, Germany experienced a sharp shift to the right. The three parties which had formed a government coalition during the previous legislature suffered big losses: the CDU/CSU and the SPD, the most important center-right and center-left parties of the traditional political system, together lost about 15 percent of their vote share and 105 seats out of about 700 members of the Bundestag, the German lower House of

[1] Number consists of "Jihadi-inspired extremists", not necessarily officially claimed by ISIL, and the ones claimed by ISIL (National Consortium for the Study of Terrorism and Responses to Terrorism 2018).

[2] See Süddeutsche Zeitung (10.07.2016) for a review of the report by the Federal Criminal Police Office of Germany.

[3] See Zeit Online (05.01.2016) for a brief summary.

Parliament. This went to the advantage of the Liberals, whose vote share increased from below 5 to 11 percent, and especially of the far right AfD (*Alternative für Deutschland*) which increased its vote share from 4.7 to 12.6 percent and, importantly, entered the Bundestag for the first time winning a total of 94 seats, 14 more than the liberals. AfD had been created just before the 2013 elections. Initially it was a Eurosceptic party concerned mainly with economic issues; it took a significant right turn after July 2015 following a clash between moderate-right and far-right members. Since then migration and external threats have become its top priorities (Cantoni et al. 2017).

This paper aims to understand what drove the shift in votes observed in the 2017 election. To what extent the salience of issues such as Islamic immigration and external threats affected the election outcome beyond concerns for economic conditions and inequality? We argue that the reasons for the rise of the AfD should also be found in the cultural shocks associated with the refugee inflow and in the terrorist attacks, as such incidents may have had an influence on German voters. For instance the terrorist attacks occurred in France (in Paris on November 13, 2015 and in Nice on July 14, 2016) could have influenced the German perception of the threat of Islamic terrorism. The inflow of refugees from Islamic countries had been singled out by the far-right such as AfD as a major threat to European identity and security. As violent incidents occurred these parties claimed to be proven right. But to what extent did this affect votes?

We argue that because of low frequency and well-known problems of self-reporting, surveys may not be able to isolate the public reaction to these violent incidents. We thus use high frequency data from Twitter and a Natural Language Processing (NLP) algorithm (Doc2Vec) that allows us to compute the distance between the content of the tweets posted by citizens and those posted by the seven main German parties. Following a growing literature in opinion mining, we surmise that the distance in the use of language captures the alignment between the public mood and the political parties. Our analysis is at the disaggregated level of German electoral constituencies: this allows us to control for constituency specific characteristics such as the number of immigrants and refugees, unemployment, inequality etc.

Language distance could change either because the language used by citizens changes, for example following a terrorist attack, or because what changes is the language of the parties' tweets. In consequence, a decreased distance could be the public moving closer to the party or vice versa. We show evidence suggestive that it is indeed the public moving closer the AfD.

Twitter users are not required to reveal where they live: we thus design a rule to attribute them to a specific electoral constituency. To validate our measure of language similarity as a proxy for the mood of voters we analyze about 31 million tweets and show that the measured similarity between these tweets and those of the seven main German political parties is positively correlated with votes and polling data. Tweets are collected whenever produced, so the data is essentially continuous time data.

Next, we use the date of the events discussed above to identify exogenous shifts in the aforementioned language similarity. Using a discontinuous growth model we compute the "jumps" in language similarity that occur in each constituency on the day of an event and that vary across constituencies: these are exogenous (that is uncorrelated with constituency characteristics) because we use random coefficients to control for constituency-specific characteristics. We observe that following such events the similarity between the text of the average constituency-level tweets and those issued by the AfD increases, while the change in similarity is lower or negative for the other parties. We then regress differences in vote shares from 2017 to 2013, at the constituency level, on exogenous shifts in language similarity. We find that exogenous components of shifts in language similarity are significant in explaining changes in vote shares.

With an estimated overall size of less than five million Twitter users, Twitter is not the most used social network in Germany. Nevertheless, we argue and discuss that the effect we measure can be seen as a lower bound for a possibly much larger effect in the overall German population.

Finally we estimate how language similarity evolves, constituency by constituency, after an exogenous shock in similarity has occurred, from the day of an event and up to the election. Because this evolution is potentially not exogenous it is interesting to investigate how language similarity evolves over time and how this evolution is correlated with the constituency characteristics we can observe: for example does a larger foreign population amplify the initial jump compared to constituencies with a lower one? We then perform placebo checks to ensure that comparable language changes are not observed during other salient public events and that the estimated effects of terrorist attacks and cultural incidents are not explained by a pre-existing movement in similarity between the text of voters and parties. Finally, we discuss the observation that voters are becoming more similar to AfD and not the other way round. We also address the possibility that non-cultural issues on the political agenda of the AfD, such as its negative stance towards the euro, are driving the increased similarity, concluding that this is unlikely to be the case.

2 Related Literature

Economic Determinants of the Growth of Populism

Our work intersects several strands of literature. We speak mostly to the area of political economy trying to explain the rise of populism in Western democracies. Although the literature on this topic grew fast in recent years,[4] the study of modern far-right political formations dates back to at least the 1990s (Betz (1994) and Kitschelt (1995)).

The rise of populism is associated to rising unemployment and increasing economic insecurity (Knigge (1998), Funke et al. (2015), Inglehart et al. (2016), and Rodrik (2018)). Colantone et al. (2018c), Colantone et al. (2018a), and Autor et al. (2016) establish a clear connection between displacements due to globalization and import competition and the rise of nationalist parties and politicians advancing extreme and protectionist views.

Some studies exploiting micro data suggest different explanations for right-wing populist support, among which protection of national identity against outsiders (for instance, Oesch (2008)). More recently, using a vast collection of administrative data, Dal Bó et al. (2018) argue that populist voters and candidates in Sweden share some characteristics of social groups affected by increased insecurity in the labor market.

This literature also highlights the interaction between economic shocks and socio-cultural attitudes, among which the decline of trust in institutions. For example, Algan et al. (2017) find that lower trust in political institutions is associated to a rise in unemployment and a fall of GDP. Similar findings are described by Dustmann, Eichengreen, et al. (2017).

Immigration and the Growth of Populism

Another reason for populist support often discussed is reaction to immigration. While the traditional literature has long identified a correlation between populist vote and anti-immigration attitudes (for instance Golder (2003)), recent work has focused on identifying the causal effects of immigration flows and disentangling heterogeneous channels. Using data from Italy, Barone et al. (2016) show that immigration moves votes to the right of the political spectrum and that the main drivers are fears of increased competition in the labor market and for welfare. Dustmann, Vasiljeva, et al. (forthcoming) study a Danish policy of random refugee allocation and find that more refugees increase votes for right-wing parties, the effect being larger in smaller and more rural municipalities and in those with higher crime rates. Competition for compositional amenities, that is the quality of local public goods and services, is another identified channel, starting from the work by Card et al. (2012). Halla et al. (2017) find that immigrant inflows in Austria induce fear of deterioration of local services and move votes towards the far-right. Hangartner et al. (2017) propose even more extreme conclusions: studying the case of Greek islands, they show that even the exposure per se to refugees increases native hostility towards immigrants and minorities. Somehow mixed findings are presented in Steinmayr (2016) for Austria and Vertier et al. (2018) for France, who show how exposure to refugee relocation can in some settings decrease support for the far-right. Notably, in some cases the perceived threat from immigration can exceed the actual threat by large amounts as found, for example, by Alesina et al. (2018) using survey data and experiments in six countries.

Germany seems a particularly suitable context in which to study different drivers of populism other than economic distress, since the country was much less affected by the Great Recession compared to many other Western countries, and fails to show sustained tension in the labor market. It is interesting to note that for the party we are focusing our research on, AfD, some scholars

[4] See Arzheimer (2019) for a comprehensive bibliography in political science

find clear differences relative to other Western European right-wing parties (Hansen et al. (2019), Goerres et al. (2018), as we shall discuss in more detail later.

Twitter, Attitudes and Politics

Our work speaks also to the literature on social media, Twitter in particular, and political behavior. Due to its easily accessible Application Programming Interface (API) and its grown popularity, data from Twitter are increasingly used to investigate a variety of topics in social sciences. Differently from surveys administrated through questionnaires, data from social media allows to "nowcast" attitudes, as noted by Curini et al. (2015), including tracing their fluctuations after the occurrence of specific events. Another advantage of using Twitter, compared to other social networks, is found in its nature of openly exchanging messages ("Tweets") as either a statement on its own or as a reply to another message. Also the ease with which common topics of discussion ("Hashtags") and the large body of information incorporated in each Tweet, including data about the author and its networks, can be retrieved from the API makes Twitter attractive to researchers. Application of Twitter data to the study of politics include, among others, investigations of the political preferences of either a population of individuals (Barberá et al. (2015), Huberty (2015), and Vaccari, Valeriani, Barberá, Bonneau, et al. (2013)) or more specific actors of public interest (An et al. (2012), Park et al. (2015), Cantoni et al. (2017), and Sterling et al. (2018)). Other topics for which Twitter provides a valuable resource are the diffusion of information (Halberstam et al. (2016), Vaccari, Valeriani, Barberá, Jost, et al. (2016), Gorodnichenko et al. (2018), and Jann et al. (2018)) and the use of networks in the context of political campaigns (Safiullah et al. (2017), Huberty (2015), and Jin et al. (2017)).

Text Analysis

In this paper we use a Natural Language Processing (NLP) algorithm - Doc2Vec - to learn a vectorial representation of the language used in Tweets. The last decade has seen a rapid evolution of applications of text analysis. For an overview, refer to Liu (2012) and Ravi et al. (2015). Scholars tend to use those techniques to either analyze the language of one group of interest with respect to some pre-defined characteristics (Curini et al. (2015), Sterling et al. (2018)) or compare languages of two or more groups of interest (Gentzkow et al. (2010), Obschonka et al. (2018)).

Language, terrorism and German politics

Our paper is also closely connected to recent work about the roots of support for the German far-right, the effects of terrorism on attitudes and partisan differences in language in social media. Among the first category, Cantoni et al. (2017) use text analysis and Twitter data to measure the persistence of right-wing ideology in Germany. Interestingly, they show the existence of a historical correlation between support for the Nazi party in the 1920s and 1930s and support for the AfD from 2015 onward. Among the second, though not in the German context, Sterling et al. (2018) compare differences in language use between liberal and conservative U.S. residents and members of the U.S. congress. They focus on values represented in the tweets of both groups, such as group loyalty and authority, and find that indeed language use reflects characteristics which are commonly associated with these groups, such as that liberals are more concerned about fairness issues. Finally, some recent work has studied public opinion in relation to the recent wave of terrorism. Giani (2017) uses the coincidence of the collection of a social survey at the time of the Charlie Hebdo shooting in Paris in 2015 and finds an increasing support for institutions engaged in

counter-terrorism, following the attack. More surprisingly, his findings show a decreasing support for reducing Muslim immigration. Similarly, Legewie (2013) exploits terror attacks in Indonesia and Madrid as natural experiments to analyze the effect of terrorist attacks on attitudes toward immigrants in European countries using Eurobarometer data. He finds evidence of cross-national variation, which is correlated with local characteristics. For example, increasing unemployment rates are positively correlated with the size of the effect of those attacks.

3 Background: AfD, Terrorism and the Refugee Crisis

The rise of AfD in German politics

AfD (*Alternative für Deutschland*, Alternative for Germany), was founded at the beginning of 2013 by an academic economist, a conservative politician and a journalist. The new party was aimed at promoting economic and fiscal conservatism and giving representation to groups of civil society which were worried by the potential risks that the European mechanisms of assistance to distressed countries in the South of the continent could pose to German public finances.

AfD first ran at the 2013 Federal Elections, barely missing the threshold of 5% of nationwide votes for winning seats in the German Parliament.[5] Still, its first-time performance was remarkable.[6] The party gained a large part of its votes in the Eastern regions of the country.

At the European elections of the following year, AfD ran again with a Eurosceptic program and gained 7.1% of nationwide votes, securing its first seats at the European Parliament. By that time, a strong tension had emerged within the party between the more centrist, fiscally conservative faction led by economist Bernd Lucke and a grown far-right leaning wing, markedly nationalistic and anti-immigration, led by an entrepreneur Frauke Petry. As AfD rose in its electoral successes, winning seats in regional elections and consolidating a constituency in the Eastern regions of the country, the national conservative faction gained importance. July 2015 is the moment which political scientists and journalists see as the crucial point in the life of AfD. A national meeting had Petry elected as party leader and Lucke leaving the party with thousands of members of the moderate wing to later create another political formation. From that moment, the party agenda shifted markedly from economic issues to criticism towards the immigration policy brought about by German Chancellor Angela Merkel and, soon after the 2015 refugee crisis overlapped with a season of Islamist terrorism in Europe, to the connection between multiculturalism and threats to security. The content of the party's messages in the social media reflect this tendency, with nationalism and Islam gradually replacing the Euro among the main topics in the AfD accounts (Cantoni et al. 2017).

Between 2015 and 2017 AfD ran several state elections, where it confirmed its growth and elected representatives in all state Parliaments, at the expenses of the more moderate center and left-wing parties, most notably Christian-Democratic CDU and Social-Democratic SPD. At the 2017 Federal election, AfD was the third party in terms of votes in the nationwide list and of seats gained in Bundestag, the German Parliament.

In recent years, after the party turned into a far-right populist formation, research in political science has studied its electoral base. The most recent cross-country evidence on populist voting has underlined the importance of "globalization losers" in the constituency of populist parties in Europe (Colantone et al. 2018b). Case studies of AfD argue that attitudes about immigration policy and German institutions (Hansen et al. 2019), as well as conservative economic policy preferences (Goerres et al. 2018) rather than economic insecurity are at the basis of support for the party. Cantoni et al. (2017), as already mentioned, bring evidence in favor of cultural channels behind populist voting in Germany: they show that the geographic distribution of votes for AfD has become correlated with that for the Nazi Party after the nationalist turn of 2015.

[5] In German electoral system, voters cast two votes: one for a constituency representative elected through a first-pass-the-most system and one for party lists.

[6] See for instance Spiegel Online (23.09.2013)

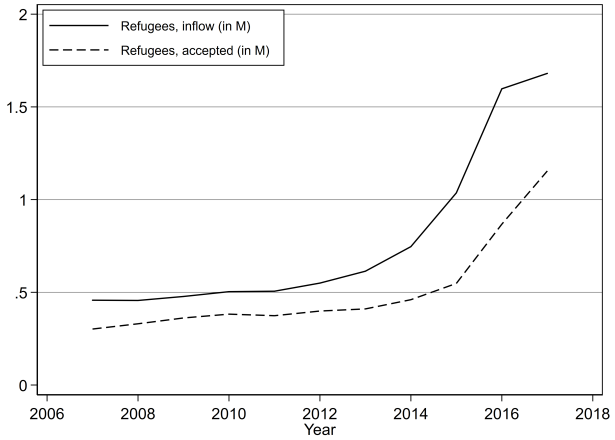


Figure 1: Refugee Inflow and Legal Acceptance

The refugee crisis and terrorism

The social and political context in which AfD rose to power was marked by a dramatic increase in immigrant flows to European countries, including Germany (see Figure 1). Major military conflicts in Syria and Sub-Saharan Africa, together with economic factors, generated large movements of people towards North and Western Europe, through the Mediterranean Sea or the Balkans. According to Eurostat, almost 4 million people applied for asylum in one of the European Union countries from 2014 to 2017. The political reactions in the Continent were not uniform. German Chancellor Angela Merkel announced in September 2015 that she would open German borders for Syrian asylum-seekers, thus inaugurating an open and welcoming policy towards refugees. The refugee policy became soon one of the crucial issues around which AfD concentrated its opposition to the Merkel government.

Parallel to the refugee crisis, starting in 2014, Western Europe was hit by an unprecedented wave of religious terrorism. That period saw the rise of ISIL and a subsequent prolonged period of conflict in the Middle East and North Africa, involving local governments and NATO troops. A number of terrorist attacks were perpetrated in some of the most important European countries, including France, Germany, Belgium and the United Kingdom.[7] Some of the perpetrators were affiliated to ISIL or to other radical Islamist organizations, most notably Al Qaeda, but in several cases the attackers were common citizens, often second-generation immigrants, Muslim or recently converted, which had rapidly radicalized. While most of the attacks were carried out with weapons, guns and bombs, "lone wolf" terrorists often resorted to common objects which were immediately available to inexperienced killers and which were readily converted in instruments of death: knives, trucks and autovehicles. Some of the deadliest attacks of that period have actually been executed with such improvized weapons. The unexpected violence of the attacks and the large number of victims, together with their relatively high frequency (from 10 to 15 between 2014 and 2016, according to different classifications) and the new profile of "next-door terrorist" which emerged, generated widespread fear and a sense of threat in the public across Europe. The political debate about security and immigration policy was inflamed and far-right and nationalist parties gave voice to fears about the threat coming from multiculturalism and failed assimilation of immigrants. In Germany, criticisms about Merkel's government and its immigration policy grew even more in light of these events.[8]

[7] The first violent attack of this upsurge of Islamist terror was the shooting which happened in January 2015 at the headquarter of magazine *Charlie Hebdo* in Paris and, few days after, in a kosher supermarket of the same city.

[8] see for instance Politico (20.12.2016)

We argue that this wave of terrorist events - as we shall explain in the following section, there was also another salient cultural incident in Cologne on New Year's eve 2015/2016, not related to a terrorist attack - opened the door to a cultural backlash against immigrants and that one of the main drivers was the perception of an external cultural threat which became abruptly more salient. In the next sections we use some of these terror events to derive measures of shocks to attitudes about culture.

4 Data

The data used throughout the project includes data from German Twitter users, statistics on municipalities and refugees provided by the Federal Statistical Office of Germany, and statistics on macro variables and electoral results at the constituency level provided by the Federal Returning Officer. We also use polling data at state level provided by infratest dimap (2018). The use of the different sources is described in the following section.

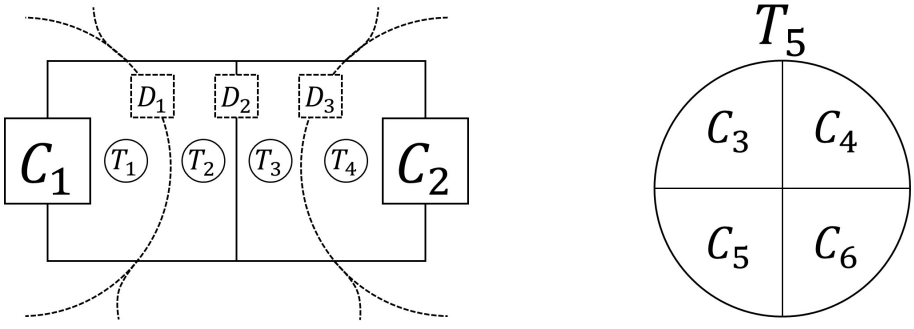
Twitter Users

The goal was to build a sample of German Twitter users which is representative of voters in German electoral constituencies. We started from a complete list of municipalities belonging to each constituency provided by the Federal Returning Officer (Der Bundeswahlleiter 2017). The first challenge was to identify where Twitter users live, i.e. the municipality where they are registered to vote. Twitter users can voluntarily choose to publish any location they wish on their profile and there is no mechanism in place to double check the provided information. Hence, using information provided by users leads to four possible outcomes. One can obtain missing addresses, correct addresses, incorrect addresses, and non-existing addresses (such as Disneyland). Excluding the latter is straightforward, but there does not exist a simple method to verify a given existing address or to replace a missing one. Most important, it is not possible to verify whether the location that an user provides is her real place of residency or not. For this reason, we constructed an algorithm that allocates users to a constituency even if they do not provide their location.

The algorithm works as follows. The 299 German electoral constituencies are drawn with the goal to equalize population across them. In consequence, electoral borders are not following any common structure, but are drawn over cities and districts.[9] By the end of 2017 there existed 401 districts and district-free cities, which correspond to the NUTS-3 classification of the European Union (see European Parliament and the Council (2003)). For a given constituency our approach first identifies the largest towns within each district that is inside the given constituency. Here we face two possible situations (as shown in Figure 2).

For a constituency (such as C_1) which contains parts of one or more districts (D_1, D_2) the largest towns in the respective districts belonging to the constituency of interest are considered (here, T_1 and T_2). Because one district (D_2) can overlap with several constituencies, the chosen towns are not necessarily the largest towns in their districts (T_3 might be larger than T_2, nevertheless T_2 is the largest town within D_2 that is still part of C_1).

[9] They also change from one election to another. The 299 corresponds to the general election of 2017 (Der Bundeswahlleiter 2017).



(a) Case 1 (b) Case 2

Figure 2: Sampling Mechanism

The second case concerns multiple constituencies (C_3 - C_6) which are entirely located within a district-free city (T_5). For instance, the city of Berlin is divided into eleven constituencies. In these cases, we merge all constituencies into one. In the end, our sample consists of 261 constituencies, either original or artificially merged, in which we chose a sample of 493 towns.

Within those towns we identify the Twitter accounts of so called landmarks. These are public or commercial accounts which can be clearly located in a given town and are followed by residents. Examples are small-scale shopping facilities, town halls, police stations, fire departments, theaters and more. Other accounts, such as sport clubs, TV stations or newspapers were not considered, because their attraction is not bound to local residents. For example, we assume that following a famous football club or a well-established newspaper is not a reliable source to infer the home town of users. Similarly, the catchment area of possible landmarks in constituencies outside of towns is much less clear than for landmarks within a town. For example, large shopping centers might attract people from relatively far away towns and using them can lead to wrong attributions to a constituency. This strategy produced a sample of 5,231 landmark Twitter accounts.

Having identified local landmarks, we used the Twitter API to fetch all the followers of those landmarks. We then restricted the sample, dropping users who followed less than three landmarks in the same constituency or landmarks in more than one constituency. This strategy produced a final sample of 178,271 Twitter users. The exact amount of active German Twitter users is unknown; different sources estimate it between 2 and 5 million users.[10] For our sample we downloaded all available tweets. Twitter limits the access to roughly the latest 3,200 tweets, but since this limit was only reached by 128 users we consider its influence on our research negligible and conclude that we have essentially all the tweets sent by our sample of users.

We follow the strategy of including Twitter users from all possible German constituencies. Importantly, we clearly state that we have no reason to believe that a representative Twitter user from our sample equals a representative German voter for a given electoral constituency. There is clearly a mechanism of self-selection in our sample. Usually, Twitter users tend to be disproportionately young, educated and from middle to high-income families.

Thus the demographic profile of Twitter users in Germany is not close to a representative supporter of a right-wing populist party, which is on average older and less educated.[11] Hence, a representative user from our sample is likely to be more moderate or liberal in political beliefs than a potential AfD voter. Therefore, we argue that although we will not be able to identify an externally valid effect for the German electorate, we will probably underestimate it. Simply formulated, we argue that if there is an effect in the population of Twitter users, then one has reason to believe that there is even a stronger effect in the German population.

Parties

The parties considered in our research are those which made it into the *Bundestag* in 2017: Alternative für Deutschland (AfD), BÜNDNIS 90/DIE GRÜNEN (GRÜNE), Christlich Demokratische Union Deutschlands (CDU), Christlich-Soziale Union in Bayern (CSU), Die Linke, Freie Demokratische Partei (FDP), and Sozial- demokratische Partei Deutschlands (SPD). For each party we considered the main national level Twitter account, excluding the personal accounts of party leaders and representatives. Table 1 shows descriptive statistics of those accounts. To overcome Twitter API limits, which allow to download only the latest 3,200 tweets, while these

[10] In 2016, Twitter announced that 12 million people are in touch with Twitter in Germany, including both registered active users and non-registered users. Considering only registered users will lead to a (much) lower number, which can only be estimated.

[11] See for instance Roth et al. (05.10.2017)

Table 1: Twitter Accounts of Major German Parties

Party	Party Account	# tweets	# Followers	Joined
AFD	@AfD	18,600	130,000	Sep-12
Bündnis 90/ Die Grünen	@Die_Gruenen	18,000	441,000	Apr-08
CDU	@CDU	16,300	274,000	Feb-09
CSU	@CSU	14,800	186,000	Feb-09
Die Linke	@dieLinke	24,500	254,000	Jun-09
FDP	@fdp	10,900	331,000	May-09
SPD	@spdde	32,200	354,000	Mar-09

Retrieved February 11, 2019. Amount of tweets includes Retweets to better reflect activity.

party accounts differently from our sample of users all have many more tweets, we web-scraped the parties' Twitter pages directly.

Electoral and Structural Data

For the years of federal elections, the Federal Returning Officer in Germany (*Bundeswahlleiter*) publishes the election results and structural variables for each electoral constituency, since electoral constituencies are not following the borders at the administrative district level (NUTS-3). Hence, data on the amount of refugees within each electoral constituency were downloaded at the administrative district level from the Federal Statistical Office for Germany and aggregated at the constituency level. For the constituencies we merged due to their size, a weighted average with respect to population was used to aggregate variables of interest.

Apart from electoral results we also use polling data at state level provided by infratest dimap (2018). These polls are the results of the so called "Sunday's Question" (*Sonntagsfrage*): more than 1,000 eligible voters are asked which party they would vote for if there were a General Election next Sunday. As indicated by the institute itself the answers to this question should be treated carefully, since they reflect the current mood of the electorate rather than a precise forecast of voting behavior. For us, however, this offers a chance to test our measured index of similarity. In the end, we are not claiming that what we measure can be used to perfectly forecast an election months or years later. What we claim is that it reflects the increased alignment of political views to a given party.

Events

For our analysis we use six events, from the end of 2015 to the end of 2016, which we identify as large shocks to public opinion. Five out of six are major terrorist attacks. Among the several events related to terrorism and violence reported in the media between 2015 and 2016, we look for a subset which satisfies some properties. First, they need to be plausibly exogenous to local conditions. Hence, we disregard very local incidents which were independent of a national context (theft or small-scale violent incidents). Events happening in other countries are particularly appropriate to this goal. Second, they need to be large shocks, affecting public opinion not only in the area where they took place (i.e. city or district), but in the whole country. Thus, we exclude some non-deadly attacks, relatively less important events. Third, we select events which plausibly affect the salience of an external cultural threat: since all the attacks in this period had Islamist extremism as alleged or clear motivation, we believe this presumption is unfortunately realistic. The five terrorist events that we consider are:

- *November 13, 2015 in Paris, France.* Three groups of terrorists perpetrated simultaneous attacks with bombs and guns on several targets in Paris, including a stadium, restaurants and a concert hall. 130 civilians were killed and other around 400 injured.
- *July 14, 2016 in Nice, France.* A young man drove a heavy truck at high speed over the crowd in a boulevard in Niece. 86 civilians were killed and around 400 were injured.
- *July 18, 2017 in Würzburg, Germany.* A young man armed with a hatchet and a knife attacked passengers of a train, injuring 5 of them.
- *July 24, 2016 in Ansbach, Germany.* A bomb was placed outside a bar in the town of Ansbach, injuring 15 people. The bombing happened only one week after the train attack in Wurzburg and ten days after the major Nice massacre.
- *December 19, 2016 in Berlin, Germany.* A young man drove a truck over the crowd in a Christmas market in the German capital city. 12 civilians were killed and around 50 were injured.

In addition to those, we include a non-terrorist event which shocked public opinion in Germany and across Europe and which generated wide political and social reactions consistent with the cultural threat framework.[12]

- *December 31 and January 1, Cologne, Germany.* During the New Year's Eve celebrations in the city centre of Cologne (and, to a smaller extent, in other German cities), several hundreds of women were subject to mass harassments and sexual assaults, often resulting in thefts. Most reports to the police identified North African and Syian young men as the perpetrators.

[12]See for instance Reuters (10.01.2016)

5 Similarity between Texts

We compute similarity between the Tweets of the parties and the tweets of each constituency by transforming the two groups of tweets into vectors through Doc2Vec. The similarity is then measured by the cosine similarity between the two vectors. Before proceeding with Doc2Vec, we pre-process tweets.

Text-preprocessing

All the words in our corpus of tweets represent the vocabulary upon which our textual analysis occurs. Since computational time is more than directly proportional to vocabulary size, text-preprocessing is the practice of reducing the initial vocabulary in such a way that relevant information is not lost but computational time is greatly reduced. We followed standard procedures in text pre-processing with different libraries in Python. We lower-cased all words and tokenize the text, i.e., we break streams of text into single words, called tokens. We use word_tokenize from the Python module NLTK. Next, we eliminated punctuation and stopwords, namely words that recur very frequently in our corpus and, thus, have little meaning. The dictionary of stopwords is the one in stopwords in NLTK. We also removed all tokens that consist only of non-alphanumeric characters; for tweets we remove emoticons, links, **@**, and # symbols. Then, we performed "stemming", which implies conflating the variant forms of a word into a common representation, the stem. For instance, the words "ate" and "eating" are both reduced to the common stem "eat". This stemming relies on pre-existing dictionaries. We used the German Stemmer in the Python module gensim. Finally, we performed collocations, namely, identifying combinations of two words that have higher probabilities of occurring together than separately. For instance, the tokens "angela" and "merkel" have higher chances of co-occuring as the bigram "angela merkel" than separately. In this case, collocations transform the two separate tokens into just one: "angela_merkel". We used BigramCollocationFinder in NLTK. The next step is to learn how to represent these words as vectors.

doc2vec

After pre-processing our tweets, we created two types of "documents": a party- and a constituency-document. A party document is the text of all the tweets that the party posted a certain day. A constituency document is the text of all the tweets that all the users in our sample located in a certain constituency posted a certain day. Since we have 750 days in our observation period (from September 4th 2015 – the day of Angela Merkel's speech – to September 24th, 2018 – the day of the election), we end up with 750 documents for each party and 720 documents for each constituency in our sample.

We relied on doc2vec (Le et al. 2014), an unsupervised deep learning algorithm that learns how to represent each document with a unique vector, and which is a generalization of word2vec. In order to understand doc2vec it is necessary to first understand how word2vec works. Word2vec (Mikolov et al. 2013) is an unsupervised deep learning algorithm that learns how to represent each word as a vector, depending on the surrounding (context) words. It takes as input a large vocabulary of words, trains a neural network language model with a single hidden layer, and produces a vector space, where each word is represented as a vector in this space. Word vectors, also called word embeddings, are positioned in the vector space such that words with similar semantic meaning are located in close proximity to one another. The model is trained using stochastic gradient descent with back propagation. When the algorithm converges, it represents words as

word embeddings, namely meaningful real-valued vectors of configurable dimension (usually, 300 dimensions).

As an extension of word2vec, (Le et al. 2014) introduced doc2vec to learn embeddings of documents. Doc2vec is an extension of word2vec that learns to capture not just individual words, but entire sentences and paragraphs. By treating each document as a word token, the same word2vec methodology is used to learn document embeddings (Bhatia et al. 2016). As in word2vec, training happens through back propagation: the algorithm is recommended to train the doc2vec model several times with unlabeled data while exchanging the input order of the documents. Each iteration of the algorithm is called an epoch, and its purpose is to increase the quality of the output vectors. This type of document embedding allows to represent texts as dense fixed-length feature vectors, taking into account their semantic and syntactic structure.

We use the Distributed Bag of Words (DBOW) model and a freely available implementation of the doc2vec algorithm included in the gensim Python module, whose implementation requires the following hyperparameters:

- Size: the dimensionality of the vector representing the document. We set it to 300.
- Window size: The maximum distance between the current and predicted word within a sentence. We set it to 15.
- Epochs: Number of iterations over the corpus to train the algorithm. We set it to 300.
- Min_count: Ignores all words with total frequency lower than this. We set it to 20.
- Subsampling: The threshold for configuring which higher-frequency words are randomly downsampled, useful range is (0, 1e-5). We set it to 1e-3.
- Negative: The number of "noise words" that should be drawn. We set it to 5.

Computing Similarity

The doc2vec gave us vectors representing each party and each constituency documents. We then measured similarity between party p and constituency c in day d as the cosine similarity between the two corresponding vectors (i.e., the normalized inner product of the two vectors):

$$\cos\theta_{abd} = \frac{\overrightarrow{a_d}\overrightarrow{b_d}}{\left\|\overrightarrow{a_d}\right\|\left\|\overrightarrow{b_d}\right\|}$$

6 Validation

The natural question arising after applying our algorithm is whether the produced similarity is indeed a valid measurement for how close public opinion is to the various national parties. For the purpose of validation we need data of public opinion for comparison. As a main check we use the electoral results in which we are most interested in, i.e. results for the 2017 federal election at constituency level. We also use results of the poll surveys provided by infratest dimap (2018) at state level to test our measure of similarity. In practice,we investigate whether the index of similarity is positively correlated with both measures.

For the election outcomes, we merge the texts we obtain from Tweets in the 30 days before the election within electoral constituencies and apply the doc2vec algorithm. The same is done for 15 days before the election as a robustness check. The reason for this merge is to produce a considerable amount of text for both parties and constituencies: not all parties issued Tweets close to the election. This may seem counter intuitive, however note that we are only using Tweets, i.e. messages written by the main party account itself, and not any other kind Twitter activity, such as Retweets or replies. For the analysis of poll data, since poll surveys are conducted at state level, we merge the texts from Tweets of all constituencies within a given state on the day of the poll.

We then use as dependent variable the change in vote share from 2013 to 2017 for all parties combined[13] and the poll results obtained and regress them on the measured similarity. We cluster standard errors on the lowest aggregate for the units of observation, i.e. electoral constituency level and state level. For the regression on poll results in levels we include party fixed effects to control for traditional variations in levels of party support. Results are presented in Table 2. We observe a positive correlation for all three estimations. Although one should not interpret the reported coefficients at this point, we conclude that our computed similarity captures the public mood across states and electoral borders to a degree that enables us to move on.

Table 2: Validation: Measured Similarity

	Δ Vote Share		**Poll**
Similarity in …			
Election − 15days	0.0982***		
	(0.0164)		
Election − 30days		0.126***	
		(0.0186)	
2015 to 2017			0.0679***
			(0.0218)
cons	-0.0207***	-0.0238***	0.102***
	(0.00381)	(0.00388)	(0.00907)
N	1122	1200	454
R^2	0.020	0.031	0.684

Notes: Δ Vote Share refers to the difference in electoral results between 2017 and 2013. In estimations of electoral outcomes (1) and (2) standard errors are clustered on constituency level. Poll refers to poll surveys on state level conducted between 2015 and 2017. In estimation of poll results (3) standard errors are clustered on state level and party fixed effects are included. Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

[13] That is, each observation is party-constituency or party-constituency-date.

7 Identification Strategy

Effect of Events

We are interested in the identification of the effect of the set of events presented in section 4 on the support for political parties at the electoral constituency level. Our analysis relies on the plausible assumption that this subset of events represents exogenous shocks whose occurrence is independent of any local characteristics. The size of the effect of a specific event, however, could differ across constituencies because of different characteristics. In other words, the degree to which a population reacts to an event is unlikely to be random and may depend on some unobservable characteristics.

Recall that we measure similarity to all parties across German electoral constituencies and over time, resulting in a panel data structure. In this panel the frequency of the data is daily. The events all happened in the evening, which means that under the assumption of no other significant event on the respective days, the change in language similarity we measure is the sudden change between the time of the event and midnight. This corresponds to the immediate reaction of the public under the influence of the event.

Exploiting the panel structure, we are able to isolate the effect of the occurred events by controlling for time invariant characteristics at the electoral constituency level using a random coefficients model. For each party we estimate separately a Discontinuous Growth Model (DGM) of measured similarity (Bliese et al. 2016) using the set of events as possible points for discontinuity. That means, we allow for a change in intercept and slope of similarity in each constituency at the day of events. For each party p we estimate

$$\begin{aligned} simil^p_{ti} = \pi^p_{0i} + \pi^p_{1i} Time_t + \pi^p_{2i} Time^2_t + \pi^p_{3i} E_t \\ + \pi^p_{4i} Reset_t + \pi^p_{5i} Reset^2_t + \epsilon^p_{ti} \end{aligned} \tag{1}$$

where index i in the coefficients represents the respective random coefficient which we use to control for time-invariant differences across constituencies. For example, $\pi^p_{1i} = \pi^p_1 + r_{1i}$ where r_{1i} represents the random coefficient indicating between-constituencies differences in time trend. The estimates we obtain are free of endogeneity caused by constituency characteristics. $simil^p_i$ is the vector which contains the measured daily similarity for party p in constituency i over all periods; $Time_t$ and $Time^2_t$ refer to the vectors of time and quadratic time trend and estimates how similarity would evolve in the absence of events; E is the matrix containing events indicator vectors and it is coded 0 prior to the event and 1 after the event; the associated parameter π^p_{3i} estimates the extent to which the intercept changes after the event. $Reset$ and $Reset^2_t$ are vectors coded 0 the day the event first occurs and it increases with each subsequent observation (day); the associated parameter π^p_{4i} indicates the degree to which the event alters the slope π^p_{1i} for the linear effect of time; π^p_{5i} indicates the extent to which the event alters the quadratic effect of time estimated with π^p_{2i}.

Summarizing, we estimate the evolution of similarity by using a quadratic time trend and allowing for random effects in changes in intercept and slope. For the structure of the variables used see Table 5 in the appendix.

For $k = 1, \ldots, 6$, the vector of coefficients $\Pi = \{\pi^p_{3ik}\}$ represents the change in similarity to all parties across electoral constituencies at the time of the six events. As we show in the results, these coefficients can be positive or negative.

Next, we pool all parties together and estimate the average effect of changes in similarity induced by the set of events on electoral outcomes.

$$\Delta vote = \alpha + \sum_{k=1}^{6} \Pi_k \, \beta_k + \upsilon \tag{2}$$

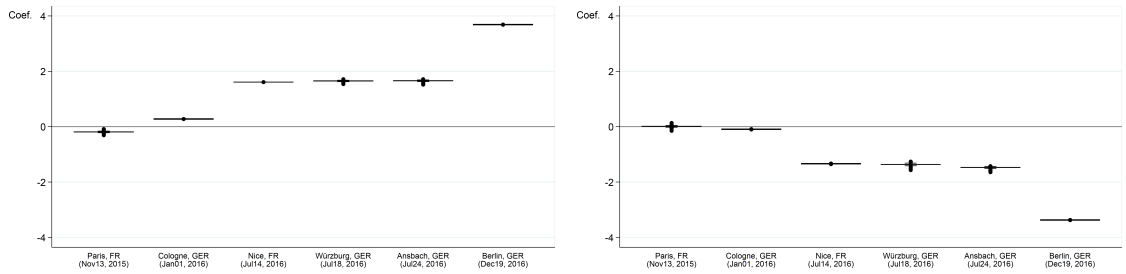
where the vector $\Delta vote$ contains the difference in vote share for for each party across constituencies between the general elections in 2017 and 2013.

Evolution over time

Support for parties evolves over time, and so does possibly the effect of the events. It may decay or increase across constituencies. Hence, we have no reason to assume that, although on the day of events we control for constituency characteristics using random coefficients, measured similarity is not affected over time by constituency characteristics. For each party the fitted values of Equation **??** represent the predicted effect of events within constituencies on a certain day. We now look at the time close to the Federal Election and investigate what characteristics shape the magnitude of this predicted effect of all our events combined. That is, we want to estimate which characteristics magnify or lower the fitted values $\widehat{simil}^{Tp}$ for a given party p at a fixed day T before the election. Our analysis in this case is correlational, since we rely on structural data at the constituency level, which do not have a consistent time dimension. We thus estimate

$$\widehat{simil}^{Tp} = \alpha + X \, \beta + \epsilon \tag{3}$$

where $X = (X_1, \ldots, X_k)$ represents a set of constituency level structural variables of interest.



(a) AfD (b) SPD

Figure 3: Jumps in Similarity on the Day of Events, AfD and SPD

8 Results

Changes in Similarity

Our results are presented in Figure 3. We see the distribution of the estimated coefficients π_i for the changes in the initial intercept for the six examined events for two parties: AfD and traditional center-left SPD. The exact model we estimated is reported in Equation 9 in the appendix. Note that, although this graph presents boxplots, in which the horizontal lines represent the median change in intercept across electoral constituencies, variation across constituencies is clearly negligible.

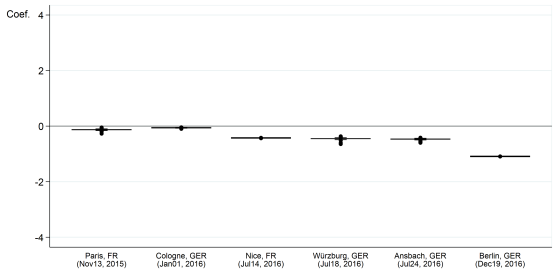
Comparing the two panels of Figure 3 we observe an opposite behavior. Similarity in language at events changes positively for the AfD and negatively for the SPD. This effect becomes stronger the more events occur, with a maximum on the day of the attack on the Berlin Christmas Market on December 19, 2016. Also, for both parties events 2 to 4 show very similar results for the change in similarity. This should not be a surprise: all of them together occur within 10 days during July 2016 and one would not expect any party to change in its rhetoric by a large amount in such a short period of time.

Results for the other parties are presented in Figure 4 and are in line with what we expected. For example, the CSU, the Bavarian counterpart of Merkel's CDU in the rest of Germany, as the traditional most right-wing party before the emergence of the AfD, also shows high jumps in language similarity, although lower compared to the AfD. This is in line with recent party history: the union of CDU and CSU was under enormous pressure during the peak of the refugee crisis around 2015. After Angela Merkel openly declared an open-border policy for asylum seekers she was challenged by high-ranked CSU officials who questioned her leadership by promoting closed borders and deportation over many months.[14] With respect to this development, seeing a similar change in similarity of language for this party and the AfD is not surprising. For the other parties the effects are negative or much lower when they are positive.

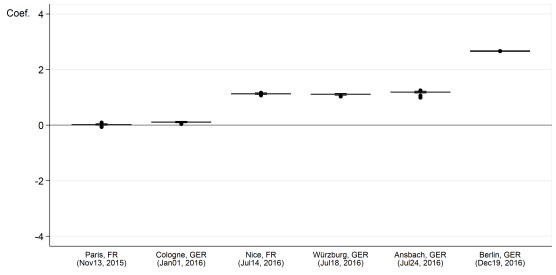
Note that being a left party, such as *Die Linke* or *Die Grünen*, does not necessarily imply negative changes in similarity. This is somewhat counter intuitive. One explanation could be that the language these parties use on the day of an event is not as distant from the language of right-wing parties, given the immediate influence of those events.

The last point motivates to start further investigation into the exact words spoken. In future versions of this work we plan to include an investigation on what the public and the parties are talking about in the hours after the examined events. We hope that this can give insights in the nature of this changes in similarity.

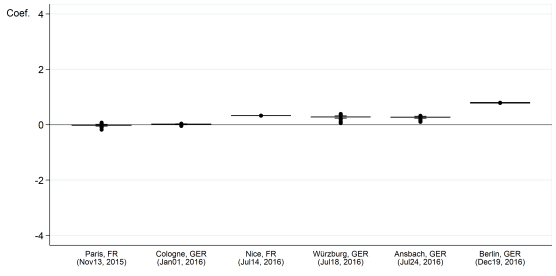
[14] See Foreign Policy (22.10.2015).



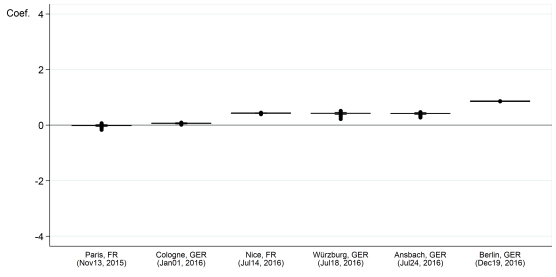
(a) CDU



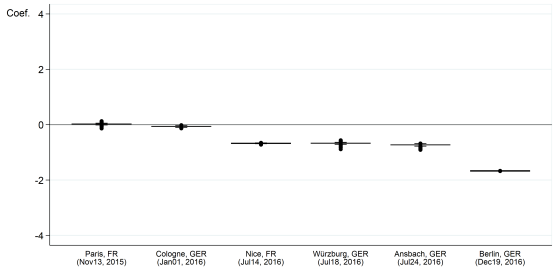
(b) CSU



(c) Die Grünen



(d) Die Linke



(e) FDP

Figure 4: Jumps in Similarity on the Day of Events, other Parties

Votes on Jumps

The last section has shown how the our terror events affect the language distance from right parties more relative to center and left parties. A natural question arising is if this change in similarity is able to explain electoral outcomes in 2017 federal election: More specifically the change in each party's 2017 vote share relative to 2013. Our research design enables us to proceed without controlling for covariates due to the use of the random coefficients model presented in Equation **??**. The reason is that we are not trying to assess what characteristics explain electoral outcomes in general. We want to investigate whether the cultural incidents and terror events contain explanatory power for electoral outcomes.

Results are presented in Table 3. We work with two different dependent variables. In the first column, we include changes in vote share for all parties. In the second we replace changes in vote share for AfD with the levels of this party's vote share. The reason is that, as explained in section 3, one can argue that the AfD of 2017 is a completely different party than the one in 2013. Hence, interpreting differences in vote shares for this party might be not as straightforward as interpreting them for all other parties. Results show a positive relationship between changes in similarity induced by the underlying events and electoral outcomes for all parties combined, with the exception of Event3 (Würzburg, July 18 2017) for which we observe a negative correlation.

Table 3: Electoral Effect: Changes in Similarity on Votes

	Δ Vote Share	**Vote Share (Δ / Lvl.)**
	(1)	(2)
Event1	0.311***	0.249***
	(0.0271)	(0.0261)
Event2	1.049***	1.343***
	(0.0577)	(0.0625)
Event3	-1.479***	-1.671***
	(0.101)	(0.101)
Event4	0.326***	0.344***
	(0.0725)	(0.0825)
Event5	0.0715	0.0714
	(0.0522)	(0.0533)
Event6	0.415***	0.477***
	(0.0245)	(0.0247)
cons	0.0403***	0.0442***
	(0.00434)	(0.00429)
N	1410	1410
R^2	0.490	0.599

Notes: Δ Vote Share refers to the difference in electoral results between 2017 and 2013. Vote Share (Δ / Lvl.) refers to the difference in electoral results between 2017 and 2013 for all parties but AfD, and to vote share in 2017 for AfD. All standard errors are clustered on constituency level and calculated using bootstrapping. Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Predicted similarity after the events

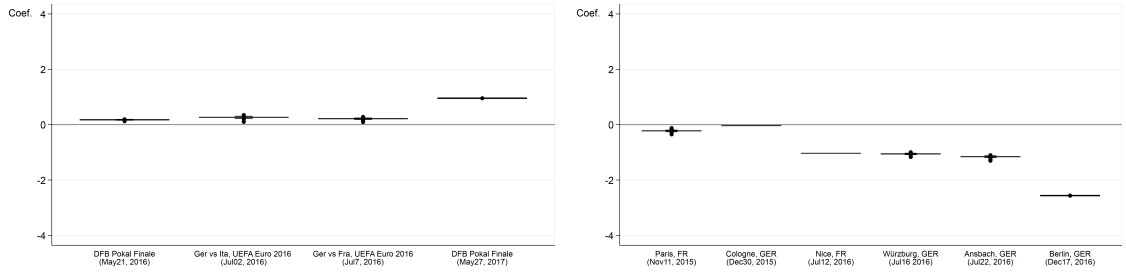
Our estimates of the constituency level effect of jumps in similarity have a causal interpretation since we have controlled for constituency characteristics using random coefficients. This, however, cannot be done for changes in similarity after the day of events. Still it is interesting to know whether constituency characteristics are correlated with an amplification or a dampen effect on similarity after an event has occurred. These results do not have a causal interpretation, they are just correlations of interest. In practice we investigate the correlation between set of control variables provided for the federal election in 2017 and fitted values of our model, i.e. the combined predicted effect of all events.

Results are presented in Table 4. We start with a set of covariates used by a post-election report by Franz et al. (2018) which investigated on the correlation connected to vote for the AfD in electoral constituencies. Column (1) is a slightly adapted version of their model of best fit. Model (2) adds refugee inflow from 2014 to 2016, model (3) adds another interaction term between East German constituencies and the inflow of refugees. The dependent variable is in all three cases the predicted similarity 15 days before the election. Refer to Table 6 in the appendix for the same calculation 30 days before the election.

Table 4: Determinants of Predicted Similarity

	Predicted Similarity (T-15d)		
	(1)	(2)	(3)
Foreigners (%)	-0.00248***	-0.00225*	-0.00200*
	(0.000961)	(0.00115)	(0.00111)
Craftsmen Firms (%)	0.00161	0.00177	0.00179
	(0.00324)	(0.00314)	(0.00321)
Age 60+ (%)	0.00405***	0.00406***	0.00430***
	(0.00152)	(0.00151)	(0.00146)
Disposable Income (Hh.)	0.00000213	0.00000215	0.00000209
	(0.00000183)	(0.00000191)	(0.00000176)
Employees in Manuf (%)	0.000859**	0.000827**	0.000845**
	(0.000375)	(0.000385)	(0.000377)
Low Eduaction (%)	0.000338	0.000367	0.000308
	(0.000960)	(0.000978)	(0.000940)
Δ Unemployment	0.0134	0.0130	0.0138
	(0.00883)	(0.00899)	(0.00885)
East	-0.0163	-0.0162	0.00115
	(0.0214)	(0.0216)	(0.0241)
Refugee inflow		0.000000320	0.000000443
		(0.000000989)	(0.000000941)
East $*$ Refugee inflow			-0.0000137*
			(0.00000826)
constant	0.150***	0.146***	0.138**
	(0.0558)	(0.0551)	(0.0546)
N	217	217	217
R^2	0.193	0.194	0.201

Notes: Predicted Similarity 15 days before the federal election. Foreigners in percent of total population. Craftsmen firms registered per 1,000 inhabitants in 2014. Disposable income of private households in 2014. Employees working in manufacturing industry subject to mandatory social security in 2016. Change in unemployment rate 2017 to 2013. Indicator variable equal to 1 for East German constituencies. Refugee inflow 2016 to 2014. Standard Errors calculated using bootstrapping. Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.



(a) Soccer Matches

(b) Two Days before the Events

Figure 5: Placebo Events

We observe a negative correlation of the share of foreigners over the whole population living in a constituency. The share of people above the age of 60 shows a positive correlation, and so does the share of people working in manufacturing. All other variables show no significant correlation apart from the interaction between east and refugee inflow. Note that this effect is not significant in the double check for 30 days before the election, so we will not interpret it.

The negative correlation with the share of foreigners within a constituency fits the common narrative which says that people who are more used to foreigners might be less likely to be heavily influenced and pushed towards a right-wing populist party. Again, one should treat such an interpretation carefully. At this point of the analysis we will not interpret the coefficients of age, since this coefficient might easily be a proxy for another determinant we cannot control for: age is very likely to be a determinant of self selection into our sample. In the end we expect people of old age to use Twitter less than people of young age. For the share of employees in the manufacturing sector we refer to the literature presented in section 2 showing that this group presents, even in absence of terrorist attacks, a likely support group for populist parties. Why this group should be more strongly influenced by our underlying events is unanswered at the current stage of research.

Robustness

Placebo Tests

Are the events we have selected really special, or are they similar to any other event not related to terror or culture? We follow two strategies. First, we estimate the same discontinuous growth model to compute changes in similarities for events which we argue are orthogonal to any political context. Here we use sport events, namely the domestic German Cup finals (*DFB Pokalfinale*) in 2016 and 2017, as well as two matches during the UEFA Euro Cup in 2016. Second, we compute the same model for two days before of our events. The results for these "Placebo Analysis" are presented in Figure 5. Both analyses do not show a comparable increase towards the AfD compared to the real events examined in the former section. Changes with respect to sport events show barely a reaction, apart from the last event. Changes for two days before the real events show no increase, rather a decrease with varying intensity.

Movements in party language

One natural concern with our method comes from the nature of the measure of language similarity that we use. We can think of similarity as an equilibrium outcome generated by the interaction between one agent, the party account, and the public. In our analysis we treat essentially the parties language on social media as exogenous and assume individuals are getting "closer" or

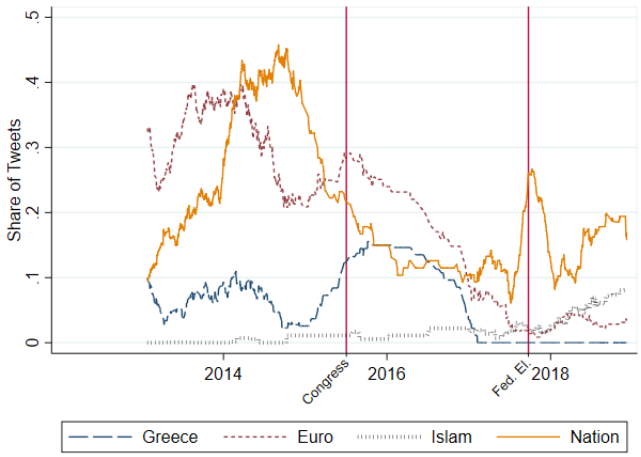


Figure 6: Frequency of specific words in AfD Tweets

"farther" from the language of different parties according to their shifting views. This assumption would be threatened if parties (AfD in particular) changed the language on their Tweets to make it more similar to that of German Twitter users, as a form of political communication strategy. We don't deny that such or similar mechanisms are part of party strategies in the social media. However, we doubt that this is a relevant factor in the context we study. First, we exploit variation in language happening the exact day of unexpected attacks. If party activism, rather than voter perceptions, was explaining our findings, this would mean that AfD accounts would be collecting reactions of Twitter users in all or most of German constituencies in real time and adapt also in real time Tweet content to the corpus of Tweets in each constituency. This hypothesis, applied to all events we consider, seems implausible.

Another possible threat would come if both parties and Twitter users tweeted about the details of events in real time, thus generating a convergence in language used. This however would not explain why the large positive shifts in similarity are observed only for AfD and are reversed for other mainstream parties.

Finally, a possibility is that voters shifted their Twitter language towards that of AfD, but for reasons that are unrelated to attitudes about immigrants and perceived threats. For instance, if the positions of AfD about the economy had become prevalent and widespread in society, rather than those on immigration and multiculturalism. We believe this is not the case. Actually, starting from the right-wing turn of the party, AfD communication in the social media has been increasingly dominated by issues of nationalism and cultural identity. The relevance of cultural issues relatively to economic ones was increasing in the period we consider, as we now show. From the full set of daily Tweets sent from the AfD account, we counted the number of times that given words appeared. In Figure 6 we plot the share of daily Tweets in which words related to "Euro", "Greece", "Islam" and "Nation" were used. We follow Cantoni et al. (2017) in the choice of keywords. We plot 3 months moving averages of the daily share of Tweets for each keyword. The graph is thus not informative about the specific events, but shows a trend. The vertical lines mark the 2015 Congress which changed the identity of AfD and the Federal Elections of September 24, 2017. In the period in between, where all our events happen, the use of word "Nation" remains constantly high, after having decreased before 2015, and increases markedly in 2017, while economic topics like "Euro" and "Greece" decrease steadily over time. "Islam" is also increasingly used between 2015 and 2016, even if to a minor extent. We are thus confident that any change in language towards that of AfD on social media is likely to be driven by cultural topics.

9 Conclusion

The rise of far-right populist parties is at the core of political and scholarly debate in Western democracies. In the European context, these parties are nationalist, protectionist and generally xenophobic. The regularly observed combination in their manifestos of opposition to free circulation of goods and people has led political scientists and economists to try to disentangle economic and cultural factors behind populist vote.

In this paper we exploit the exogenous timing of large events of terrorism and crime to isolate the effects of attitudes towards a party with a strong anti-immigrants platform. Using an algorithm based on followers networks to assign Twitter users to geographic constituencies and Natural Language Processing techniques, we show that unexpected terrorist attacks and large crime news move the language of peoples' tweets closer to that of the tweets of the main anti-immigration and nationally conservative German right-wing party. We argue that terrorist attacks and large-scale crimes committed by immigrants constituted information shocks which dramatically increased the salience of cultural differences and elicit perceptions of threats and hostility from a different religion. The effect was to move the political leaning of Twitter users towards the party which emphasized threats from immigration and multiculturalism the most.

We find that in the same day a major terrorist or culturally charged event occurs, the average electoral constituency observes a significant shift in the language of its Tweets towards the language of far-right AfD and opposite shifts from left-wing SPD and centrist CDU, the main government party. We perform placebo experiments to show that these effects are not the result of generally salient events unrelated to politics or to pre-existing trends in similarity which would artificially identify movements. Then, we show that those *party-specific* language jumps associated to events are significant predictors of vote shifts across Federal Elections. Finally, we show that predicted language similarity to AfD after the events and close to elections is associated to demographic and economic variables which are found to be correlates of AfD vote, suggesting that the context may matter in the way these shocks propagate and transform into votes.

At this stage of our research, we are cautious with respect to the interpretation of our findings. However, we emphasize that our evidence is overall consistent with the channel that we hypothesize. Specifically, jumps in language following the events we consider are unlikely to be driven by discourse over economic issues, which was being considerably downplayed in the social media of AfD during the period studied. Our results are an important first step for understanding cultural backlash as a channel of transmission of radical right support, which has been traditionally difficult to study and which we can address in the appropriate way using new data and techniques. Future work should study in more detail some of the aspects outlined here. In particular, we plan to "unpack" language similarity and study how do voters become more similar to populists.

References

Alesina, A., Miano, A., and Stantcheva, S. (2018), Immigration and Redistribution. *NBER Working Paper Series*.

Algan, Y., Guriev, S., Papaioannou, E., and Passari, E. (2017), The European Trust Crisis and the Rise of Populism. *Brookings Papers on Economic Activity*. 2017(2), 309–400.

An, J., Cha, M., Gummadi, K. P., Crowcroft, J., and Quercia, D. (2012), Visualizing Media Bias through Twitter.

Arzheimer, K. (2019), *The Eclectic, Erratic Bibliography on the Extreme Right in Western Europe*.

Autor, D., Dorn, D., Hanson, G., and Majlesi, K. (2016), Importing Political Polarization? *Working paper*.

Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A., and Bonneau, R. (2015), Tweeting From Left to Right: Is Online Political Communication More Than an Echo Chamber? *Psychological science*. 26(10), 1531–1542.

Barone, G., D'Ignazio, A., Blasio, G. de, and Naticchioni, P. (2016), Mr. Rossi, Mr. Hu and politics. The role of immigration in shaping natives' voting behavior. *Journal of Public Economics*. 136, 1–13.

Betz, H.-G. (1994), *Radical Right-Wing Populism in Western Europe*. New York: Palgrave Macmillan US and Palgrave Macmillan.

Bhatia, S., Lau, J. H., and Baldwin, T. (2016), Automatic Labelling of Topics with Neural Embeddings. *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, pp. 953–963.

Bliese, P. D. and Lang, J. W. B. (2016), Understanding Relative and Absolute Change in Discontinuous Growth Models. *Organizational Research Methods*. 19(4), 562–592.

Cantoni, D., Hagemeister, F., and Westcott, M. (2017), Persistence and Activation of Right-Wing Political Ideology. *Munich Discussion Paper*.

Card, D., Dustmann, C., and Preston, I. (2012), Immigration, Wages, and Compositional Amenities. *Journal of the European Economic Association*. 10(1), 78–119.

Colantone, I. and Stanig, P. (2018a), Global Competition and Brexit. *American Political Science Review*. 112(02), 201–218.

— (2018b), Replication Data for: The Trade Origins of Economic Nationalism: Import Competition and Voting Behavior in Western Europe. *American Journal of Political Science*. (62), 936–953.

— (2018c), The trade origins of economic nationalism: Import competition and voting behavior in Western Europe. *American Journal of Political Science*. 62(4), 936–953.

Curini, L., Iacus, S., and Canova, L. (2015), Measuring Idiosyncratic Happiness Through the Analysis of Twitter: An Application to the Italian Case. *Social Indicators Research*. 121(2), 525–542.

Dal Bó, E., Finan, F., Folke, O., Persson, T., and Rickne, J. (2018), Economic Losers and Political Winners: Sweden's Radical Right.

Der Bundeswahlleiter (2017), *Wahlkreise und zugeordnete Gemeinden bei der Wahl zum 19. Deutschen Bundestag*. Wiesbaden.

Dustmann, C., Eichengreen, B., Otten, S., Sapir, A., Tabellini, G. E., and Zoega, G. (2017), *Europe's trust deficit: Causes and remedies*. Vol. 1. Monitoring international integration. London: CEPR Press.

Dustmann, C., Vasiljeva, K., and Damm, A. P. (forthcoming), Refugee migration and electoral outcomes. *Review of Economic Studies*.

European Parliament and the Council (2003), *Regulation (EC) No 1059/2003 of the European Parliament and of the Council of 26 May 2003 on the establishment of a common classification of territorial units for statistics (NUTS)*.

Foreign Policy (22.10.2015), Nothing Can Take Down Angela Merkel — Except 800,000 Refugees.

Franz, C., Fratzscher, M., and Kritikos, A. S. (2018), *German right-wing party AfD finds more support in rural areas with aging populations*.

Funke, M., Schularick, M., and Trebesch, C. (2015), Going to Extremes: Politics after Financial Crisis, 1870-2014.

Gentzkow, M. and Shapiro, J. M. (2010), What Drives Media Slant? Evidence From U.S. Daily Newspapers. *Econometrica*. 78(1), 35–71.

Giani, M. (2017), Don't Ban the Muslims: Public Reaction to Terror Attacks.

Goerres, A., Spies, D. C., and Kumlin, S. (2018), The Electoral Supporter Base of the Alternative for Germany. *Swiss Political Science Review*. (24), 246–269.

Golder, M. (2003), Explaining Variation In The Success Of Extreme Right Parties In Western Europe. *Comparative Political Studies*. 36(4), 432–466.

Gorodnichenko, Y., Pham, T., and Talavera, O. (2018), *Social Media, Sentiment and Public Opinions: Evidence from #Brexit and #USElection*. Cambridge, MA: National Bureau of Economic Research.

Halberstam, Y. and Knight, B. (2016), Homophily, group size, and the diffusion of political information in social networks: Evidence from Twitter. *Journal of Public Economics*. 143, 73–88.

Halla, M., Wagner, A. F., and Zweimüller, J. (2017), Immigration and voting for the far right. *Journal of the European Economic Association*. 15(6), 1341–1385.

Hangartner, D., Dinas, E., Marbach, M., Matakos, K., and Xefteris, D. (2017), Does Exposure to the Refugee Crisis Make Natives More Hostile? *American Political Science Review*, 1–14.

Hansen, M. A. and Olsen, J. (2019), Flesh of the Same Flesh: A Study of Voters for the Alternative for Germany (AfD) in the 2017 Federal Election. *German Politics*. (28), 1–19.

Huberty, M. (2015), Can we vote with our tweet? On the perennial difficulty of election forecasting with social media. *International Journal of Forecasting*. 31(3), 992–1007.

infratest dimap (2018), *Umfragen & Analysen: Bundesländer*.

Inglehart, R. F. and Norris, P. (2016), Trump, Brexit, and the Rise of Populism: Economic Have-Nots and Cultural Backlash. *HKS Faculty Research Working Paper Series*.

Jann, O. and Schottmüller, C. (2018), Why Echo Chambers are Useful. *University of Oxford, Department of Economics, Discussion Paper Series*.

Jin, Z., Juan, C., Han, G., Yongdong, Z., Yu, W., and Jiebo, L. (2017), Detection and Analysis of 2016 US Presidential Election Related Rumors on Twitter. *Social, Cultural, and Behavioral Modeling*. Ed. by D. Lee, Y.-R. Lin, N. Osgood, and R. Thomson. Cham: Springer International Publishing, pp. 14–24.

Kitschelt, H. (1995), *The radical right in Western Europe: A comparative analysis*. Ann Arbor: Univ. of Michigan Press.

Knigge, P. (1998), The ecological correlates of right-wing extremism in Western Europe. *European Journal of Political Research*. 34(2), 249–279.

Le, Q. and Mikolov, T. (2014), Distributed representations of sentences and documents. *International conference on machine learning*, pp. 1188–1196.

Legewie, J. (2013), Terrorist Events and Attitudes toward Immigrants: A Natural Experiment. *American Journal of Sociology*. 118(5), 1199–1245.

Liu, B. (2012), Sentiment Analysis and Opinion Mining. *Synthesis Lectures on Human Language Technologies*. 5(1), 1–167.

Mikolov, T., Chen, K., Corrado, G., and Dean, J. (2013), Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

National Consortium for the Study of Terrorism and Responses to Terrorism (2018), *Global Terrorism Database*.

Obschonka, M. and Fisch, C. (2018), Entrepreneurial personalities in political leadership. *Small Business Economics*. 50(4), 851–869.

Oesch, D. (2008), Explaining Workers' Support for Right-Wing Populist Parties in Western Europe: Evidence from Austria, Belgium, France, Norway, and Switzerland. *International Political Science Review*. 29(3), 349–373.

Park, S., Lee, J., Ryu, S., and Hahn, K. S. (2015), The Network of Celebrity Politics. *The ANNALS of the American Academy of Political and Social Science*. 659(1), 246–258.

Politico (20.12.2016), Berlin attack exposes Angela Merkel's right flank.

Ravi, K. and Ravi, V. (2015), A survey on opinion mining and sentiment analysis: Tasks, approaches and applications. *Knowledge-Based Systems*. 89, 14–46.

Reuters (10.01.2016), Growing scale of Cologne attacks stokes German debate on migrants.

Rodrik, D. (2018), Populism and the Economics of Globalization. *Journal of International Business Policy*. 1(1-2), 12–33.

Roth, A. and Wolff, G. B. (5.10.2017), *What has driven the votes for Germany's right-wing Alternative für Deutschland?*

Safiullah, M., Pathak, P., Singh, S., and Anshul, A. (2017), Social media as an upcoming tool for political marketing effectiveness. *Asia Pacific Management Review*. 22(1), 10–15.

Spiegel Online (23.09.2013), German Euroskeptics' Breakthrough Moment: Next Stop, Brussels?

Steinmayr, A. (2016), Exposure to Refugees and Voting for the Far-Right: (Unexpected) Results from Austria. *IZA Discussion Paper*.

Sterling, J. and Jost, J. T. (2018), Moral discourse in the Twitterverse: Effects of ideology and political sophistication on language use among U.S. citizens and members of Congress. *Journal of Language and Politics*. 17(2), 195–221.

Süddeutsche Zeitung (10.07.2016), 1200 Frauen wurden Opfer von Silvester-Gewalt.

Vaccari, C., Valeriani, A., Barberá, P., Bonneau, R., Jost, J. T., Nagler, J., and Tucker, J. (2013), Social media and political communication. A survey of Twitter users during the 2013 Italian general election. *Rivista italiana di scienza politica*.

Vaccari, C., Valeriani, A., Barberá, P., Jost, J. T., Nagler, J., and Tucker, J. A. (2016), Of Echo Chambers and Contrarian Clubs: Exposure to Political Disagreement Among German and Italian Users of Twitter. *Social Media + Society*. 2(3), 205630511666422.

Vertier, P. and Viskanic, M. (2018), Dismantling the "Jungle": Migrant Relocation and Extreme Voting in France. *CESifo Working Papers*.

Zeit Online (5.01.2016), Polizei gibt Fehler in Silvesternacht zu.

Appendix A

Discontinuous Growth Model

Variables considered so far:

1. $Time$: The first variable represents the linear time trend found in a typical growth model.
2. $Time^2$: Similar to before with a quadratic time trend.
3. E: The second change variable (Event) is coded 0 prior to the event and 1 after the event. The parameter estimate associated with this variable (π_{ki}) determines the degree to which the intercept was altered after the event.
4. $Reset$: The third change variable (Recovery) is coded 0 at the period of which the event first occurs and increases with each subsequent period. It functions like a new starting time trend. The parameter estimate γ_{ki} indicates the degree to which the event alters the slope (π_{1i}) for the time variable.
5. $Reset^2$: Similar to before with a quadratic change variable.

For analyzing multiple events, we simply introduce multiple variables for events and changes. Consider the following Table 5 as overview on the variables.

Table 5: Coding of Time Variables - Multiple Events

$Time$	$Time^2$	E_1	E_2	Rec_1	Rec_1^2	Rec_2	Rec_2^2
0	0	0	0	0	0	0	0
1	1	0	0	0	0	0	0
2	4	0	0	0	0	0	0
3	9	0	0	0	0	0	0
4	16	0	0	0	0	0	0
5	25	1	0	0	0	0	0
6	36	1	0	1	1	0	0
7	49	1	0	2	4	0	0
8	64	1	0	3	9	0	0
9	81	1	0	4	16	0	0
10	100	0	1	0	0	0	0
11	121	0	1	0	0	1	1
12	144	0	1	0	0	2	4
13	169	0	1	0	0	3	9
14	196	0	1	0	0	4	16
...	...	...	...	...	...	...	...

Note that here we have the first event occurring in period 5, the second event occurring in period 10.

In our results we estimate the discontinuous growth model introduced in section 7 with a linear and quadratic time trend, the set of events as changes of intercept and a linear and quadratic change in slope. We include random coefficients. The estimated equation for measured similarity $simil_{i,t}^p$, for a given party p in constituency i at time t is in this case:

$$\begin{aligned} simil_{ti}^p = {} & \pi_{0i}^p + \pi_1^p Time_t + \pi_2^p Time_t^2 + \pi_3^p E_t \\ & + \pi_4^p Reset_t + \pi_5^p Reset_t^2 + \epsilon_{ti}^p \\ & + r_{1i}^p Time_t + r_{2i}^p Time_t^2 + r_{3i}^p E_t + r_{4i}^p Reset_t + r_{5i}^p Reset_t^2 \end{aligned}$$

Predicted Similarity

Table 6: Determinants of Predicted Similarity (30 Days)

	Predicted Similarity (T-30d)		
Foreigners (%)	-0.00245***	-0.00221**	-0.00196*
	(0.000934)	(0.00107)	(0.00111)
Craftsmen Firms (%)	0.00188	0.00205	0.00208
	(0.00317)	(0.00310)	(0.00317)
Age 60+ (%)	0.00401***	0.00402***	0.00426***
	(0.00145)	(0.00148)	(0.00152)
Disposable Income (Hh.)	0.00000195	0.00000197	0.00000191
	(0.00000178)	(0.00000183)	(0.00000187)
Employees in Manuf (%)	0.000874**	0.000842**	0.000859**
	(0.000361)	(0.000376)	(0.000364)
Low Eduaction (%)	0.000278	0.000308	0.000250
	(0.000956)	(0.000957)	(0.000918)
Δ Unemployment	0.0137	0.0133	0.0141
	(0.00865)	(0.00872)	(0.00856)
East	-0.0161	-0.0161	0.00108
	(0.0216)	(0.0207)	(0.0223)
Refugee inflow		0.000000331	0.000000453
		(0.000000889)	(0.000000944)
East $*$ Refugee inflow			-0.0000135
			(0.00000848)
Constant	0.151***	0.147***	0.140**
	(0.0552)	(0.0569)	(0.0547)
Observations	217	217	217
R^2	0.200	0.201	0.208

Notes: Δ Predicted Similarity 30 days before the federal election. Foreigners in percent of total population. Craftsmen firms registered per 1,000 inhabitants in 2014. Disposable income of private households in 2014. Employees working in manufacturing industry subject to mandatory social security in 2016. Change in unemployment rate 2017 to 2013. Indicator variable equal to 1 for East German constituencies. Refugee inflow 2016 to 2014. Standard Errors calculated using bootstrapping. Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.