Fanning the Flames of Hate: Social Media and Hate Crime*

Karsten Müller† and Carlo Schwarz‡

June 5, 2020

Abstract

This paper investigates the link between social media and hate crime. We show that anti-refugee sentiment on Facebook predicts crimes against refugees in otherwise similar municipalities with higher social media usage. To establish causality, we exploit exogenous variation in the timing of major Facebook and internet outages. Consistent with a role for "echo chambers", we find that right-wing social media posts contain narrower and more loaded content than news reports. Our results suggest that social media can act as a propagation mechanism for violent crimes by enabling the spread of extreme viewpoints.

JEL *Classification*: D74, J15, Z10, D72, O35.
Keywords: social media, hate crime, minorities, Germany, AfD

*We are grateful to Sascha Becker, Christopher Blattman, Leonardo Bursztyn, Mirko Draca, Ruben Enikolopov, Thiemo Fetzer, Evan Fradkin, Matthew Gentzkow, Andy Guess, Vardges Levonyan, Atif Mian, Magne Mogstad, Sharun Mukand, Imran Rasul, Hans-Joachim Voth, Fabian Waldinger, Noam Yuchtman, and seminar participants at the NBER Summer Institute, University of Chicago, EEA Conference 2018, Transatlantic Doctoral Conference (LBS), Oxford Internet Institute, Geneva Academy of Humanitarian Law, Bruneck Political Economy Workshop, Leverhulme Causality Conference at the University of Warwick, Spring Meeting of Young Economists 2019, the Royal Economic Society 2019, and the UNHCR Conference on Forced Displacement for their helpful suggestions. We would also like to thank the Amadeu Antonio Stiftung for sharing their data on refugee attacks with us. Müller was supported by a Doctoral Training Centre scholarship granted by the Economic and Social Research Council [grant number 1500313]. Schwarz was supported by a Doctoral Scholarship from the Leverhulme Trust as part of the *Bridges* program.

†Princeton University, The Julis-Rabinowitz Center for Public Policy and Finance, karstenm@princeton.edu
‡University of Warwick, Department of Economics, Centre for Competitive Advantage in the Global Economy (CAGE), c.r.schwarz@warwick.ac.uk

1 Introduction

Social media has come under increasing scrutiny in recent years. In the wake of the 2016 presidential election in the United States, for example, relatively recent phenomena such as fake news, social media echo chambers, and bot farms have been subjects of widespread media coverage and public discourse (e.g New York Times, 2016, 2017a). The role of online hate speech in particular has been at the center of an intense and polarized debate. Despite public interest and calls for policy action, there is little empirical evidence on how hateful social media content translates into real-life behavior.

In this paper, we investigate the role of social media in the propagation of hate crimes. Previous research has shown that traditional media can play a role in violent outbursts or ethnic hatred (e.g Yanagizawa-Drott, 2014; Adena et al., 2015; DellaVigna et al., 2014). In contrast to traditional media, social media platforms allow users to easily self-select into niche topics and extreme viewpoints. This preferential selection may limit the spectrum of information people absorb and create "echo chambers" (Sunstein, 2009, 2017), which reinforce similar ideas (see e.g. Bessi et al., 2015; Del Vicario et al., 2016; Schmidt et al., 2017). Social media has also become a widely-consumed news source, particularly for young people: in Germany, for example, social media is among the main news sources of 18 to 25 year olds (Hölig and Hasebrink, 2016). In the US, around half of all adults use social media to get news and two thirds of Facebook users use it as a news source (Pew Research Center, 2018). This suggests that social media could be particularly effective in propagating hateful sentiments.

We study the link between anti-refugee sentiment on Facebook and hate crimes against refugees in Germany. The German setting is motivated by the influx of around one million refugees into the country between 2015 and 2016 (BAMF, 2016), which was accompanied by frequent violent crimes committed against them (see, for example, recent video coverage by New York Times, 2017b). Between January 2015 and early 2017 alone, the non-profit organization "Amadeu Antonio Stiftung" recorded around 3,300 anti-refugee incidents, including over 750 cases of arson or outright assault.

We posit that social media can reinforce anti-refugee sentiments, which may push some potential perpetrators over the edge to carry out violent acts. Our empirical strategy exploits differences in Facebook usage at the municipal level and weekly variation in anti-refugee sentiment on social media. We create a novel measure for the salience of anti-refugee hate speech on social media based on the Facebook page of the "Alternative für Deutschland" (Alternative for

Germany, AfD hereafter), a relatively new right-wing party that became the third-strongest faction in the German parliament following the 2017 federal election. The AfD has positioned itself as an anti-refugee and anti-immigration party. With more than 300,000 followers, 175,000 posts, 290,000 comments, and 500,000 likes (as of early 2017), their Facebook page has a broader reach than that of any other German party.[1]

This widespread reach makes the AfD's Facebook page uniquely suited to measure anti-refugee sentiment on social media. In contrast to established political parties like Angela Merkel's Christian Democratic Union (CDU) or the German Social Democrats (SPD), the AfD allows users to directly post messages on its Facebook wall. The AfD is also the only party that does not explicitly outline rules of conduct, e.g. by threatening to remove racist, discriminating, or otherwise hateful comments. We show that the content on the AfD page is consistently more focused on refugees than that of traditional news reports and frequently contains loaded terms that civil rights groups have identified as "hate speech". These detailed data also allow us to construct a measure of each town's exposure to Germany-wide anti-refugee sentiment using the share of the population that is active on the AfD Facebook page.

Using fixed effects panel regressions, we find that—during periods of high salience of refugees on right-wing social media—anti-refugee hate crimes increase in areas with higher Facebook usage. This correlation is especially pronounced for violent incidents such as assault. Controlling for a large vector of municipality characteristics, interacted with our salience measure, makes little difference for the magnitude and statistical significance of these estimates.

A concern is that our measures of exposure to right-wing social media may be correlated with unobserved municipal characteristics that explain disproportionate increases in hate crimes during times of high anti-refugee sentiment. To narrow down the social media transmission channel, we provide quasi-experimental evidence using the exact timing of country-wide Facebook outages and local internet disruptions, which reduce the number of social media posts.

To begin, we study large, Germany-wide Facebook outages resulting from programming or server problems at the platform. These outages disrupt users' exposure to this particular social media platform without affecting other online channels. We find that Facebook disruptions reduce local hate crimes, particularly in areas with many AfD users. Further, during Facebook outages, higher anti-refugee sentiment is not associated with a differential increase in hate crimes in areas with high Facebook usage. These results suggest that social media might play a propagating role in translating online content into offline violence.

[1] We provide a short history of the AfD in Appendix A in the online appendix.

We also exploit the precise timing of hundreds of local internet disruptions as a source of granular exogenous variation in access to social media. These local disruptions reduce a particular town's exposure to social media content while leaving Germany-wide refugee salience unaffected. Notably, the frequency of internet disruptions is geographically dispersed and largely unrelated to observable local characteristics, including AfD likes on Facebook.

We find that, while hate crimes increase in periods of higher refugee salience, this correlation disappears for municipalities experiencing an internet outage. Quantitatively, a typical internet disruption fully mediates the link between social media and hate crime. Further, once we take into account social media transmission, these internet outages themselves are no longer associated with anti-refugee incidents, nor are their interactions with local internet usage or mobile internet access. These results point to social media as propagation mechanism rather than other online channels. It also makes it unlikely that we are capturing a "displacement effect" arising from potential perpetrators fixing their internet access.

We also analyze how other salient news events mediate the link of anti-refugee Facebook posts with the number of violent incidents, building on Eisensee and Strömberg (2007) and Durante and Zhuravskaya (2018). Specifically, we look at the European Soccer Championship, Brexit, and Donald Trump's presidential election, all of which crowded out the salience of refugees. Similar to our outage results, social media exposure has a significantly more muted relationship with hate crimes during these events. The link we uncover appears to be specific to anti-refugee sentiment: other posts on the AfD Facebook page, e.g. those related to Muslims or the European Union, do not have the same predictive power for anti-refugee hate crimes. Consistent with the hypothesis that social networks can act as transmission channel, the correlation with hate crime is larger in regions where AfD users show higher Facebook engagement.

When interpreting our results, we do not claim that social media itself causes crimes against refugees out of thin air. Rather, our argument is that social media can act as a propagating mechanism for hateful sentiments that likely have many fundamental sources. We find evidence for two potential channels. First, our results are driven by refugee attacks committed by groups of perpetrators. This suggests that social media may motivate collective action, consistent with existing evidence on other political outcomes such as protests (e.g. Enikolopov et al., 2016). Second, we find evidence for a spillover channel. Hate crimes are considerably more common in weeks when neighboring towns also experience them, and this is particularly true for towns with many right-wing social media users when anti-refugee sentiment is elevated. In contrast, we find little evidence that social media provides useful information to perpetrators. Our results are also

unlikely to be explained by persuasion effects, because we focus on high-frequency variation.

Related literature. Our work provides evidence that social media may have effects on real-life outcomes, as measured by hate crimes. We build on existing work on media exposure and persuasion (see e.g. DellaVigna and Gentzkow, 2010; DellaVigna and Ferrara, 2015). In addition to work on traditional media and violence cited above, Dahl and DellaVigna (2009) show that—in contrast to experimental settings—violent movies decrease violent crime in the field due to displacement effects. Television has also been associated with short-lived outbursts of domestic violence (Card and Dahl, 2011). In other research, Bhuller et al. (2013) demonstrate that exposure to pornographic material on the internet is linked to increased sex crime. Bursztyn et al. (2017) find that media coverage of close elections increases voter turnout, while Gavazza et al. (2018) show that broadband diffusion decreased voter turnout in the United Kingdom (see also Gentzkow, 2006; Manacorda and Tesei, 2016). Enikolopov et al. (2016) find that social media exposure spurs protest participation in Russia by reducing coordination costs.

We contribute to this literature by investigating the role of social media in stirring up violence. Previous research has documented the prevalence of online hate speech (Oksanen et al., 2014). Other work has shown that Google search data can be used to measure racial animus (Stephens-Davidowitz, 2014). In complementary work, we study the effect of Twitter usage on anti-minority sentiments in the United States (Müller and Schwarz, 2018). Bursztyn et al. (2019) study the effect of social media on xenophobia in Russia. In contrast to these papers, we focus on the short-run impact of social media posts, rather than long-run effects that may work through persuasion or changes in social norms.

Our paper also builds on research about the polarization of citizens (e.g Fiorina and Abrams, 2008). There is no consensus on whether social media increases or decreases polarization: some authors argue that social media are divisive (Pariser, 2011; Gabler, 2016), while others find that polarization *decreases* with social media usage (Barberá, 2014; Boxell et al., 2017). Our work suggests that—independent of whether social media affects overall polarization or not—social media content can be associated with violent crimes.

We also build on the literature on culture and violence. Summarizing a vast body of research, Alesina and La Ferrara (2005) find that cultural and and religious fragmentation predict the likelihood of civil war across countries. Voigtlander and Voth (2012) show that anti-Semitic violence in Germany is highly persistent: pogroms during the era of the Black Death predict pogroms in the 1920s, Jewish deportations, and synagogue attacks during the rise of the Nazi party. Similarly, Jha (2013) shows that medieval interethnic complementarities in

trade decrease the likelihood of modern Hindu-Muslim riots. These papers, however, are largely silent on the existence of volatile, short-lived bursts of sentiment leading to violent incidents. As such, our work is also related to Fouka and Voth (2013), who show that monthly variation in public acrimony between Greek and German politicians during the Greek debt crisis affected German car purchases particularly in areas of Greece where German troops committed war crimes during World War II. Our results also align with the findings of Colussi et al. (2016), who show that a higher salience of minority groups increases the likelihood of hate crimes.

While traditional media such as television are regulated in most countries, legislators are now beginning to address social media. Our work is thus particularly topical in light of the political discussions in many countries about anti-hate speech laws and censoring hate speech on social media. The German parliament, for example, passed an anti online hate speech law ("Netzwerkdurchsetzungsgesetz") on June 30, 2017, which threatens providers of online platforms such as Facebook with fines up to EUR 50 million for failing to delete "criminal" content that is "obviously unlawful". The controversial law was the initiative of German Minister of Justice Heiko Maas, who lamented social media platforms' unwillingness to address "online hate crime".[2] The European Union has issued independent guidelines calling on social media companies to remove illegal hate speech as well. In the United Kingdom, the Crown Prosecution Service plans to increase prosecution of online hate crimes (The Guardian, 2017; BBC, 2017). Our paper serves as a first attempt to address this important topic empirically.

The paper proceeds as follows. In Section 2 we introduce the data used in our empirical analysis. Section 3 presents the results. Section 4 concludes.

2 Data

We construct a dataset on social media activity and anti-refugee hate crimes in Germany. In total, we combine data from 12 different sources which we describe in more detail in the following subsections: (1) Municipal-level data on anti-refugee hate crimes; (2) Facebook data on posts, likes, and comments on the AfD page; (3) hand-collected municipal-level data on Facebook user locations; (4) municipal-level data on internet outages; (5) a hand-coded dataset on major weekly Facebook outages; (6) municipal- and county-level socioeconomic data from the German Statistical Office; (7) municipal-level voting data; (8) county-level data on broadband access; (9) municipal-level data on newspaper sales; (10) data on the content of reporting about refugees

[2]See, for example, the official statement of the German parliament on bundestag.de.

from Nexis; (11) city-level data on neo-Nazi murders and historical anti-Semitism; and (12) weekly Google search data on major news events in our sample. The final panel dataset covers 4,466 German municipalities for the 111 weeks from 1st January 2015 to 13th February 2017. Summary statistics for the main variables of interest can be found in Table 1 and Table A.3. The online appendix provides a comprehensive overview of the data sources and variable definitions (see Table A.4).

Table 1: Summary Statistics for Main Variables

	Level	Obs	Mean	SD	Min.	Max.
Refugee Attacks						
Refugee attacks	Muni.-Week	495,726	0.007	0.099	0	8
Arson attacks	Muni.-Week	495,726	0.000	0.022	0	2
Other property damage	Muni.-Week	495,726	0.004	0.076	0	8
Assaults	Muni.-Week	495,726	0.001	0.035	0	3
Protests	Muni.-Week	495,726	0.001	0.030	0	5
Social Media Data						
AfD users/Pop.†	Municipality	495,726	0.301	0.286	0	8
Refugee posts	Week	495,726	84	61	2	259
Posts/AfD users	Municipality	395,493	0.554	3.882	0	118
Comments/AfD users	Municipality	395,493	1.1	7.3	0	270
Likes/AfD users	Municipality	395,493	1.8	12.3	0	370
Auxiliary Variables						
$I_{Internet\ outage}$	Muni.-Week	495,726	0.001	0.025	0.000	1.000
$I_{Facebook\ outage}$	Municipality	495,726	0.072	0.259	0.000	1.000
Baseline Controls						
Ln(Population (2015))	Municipality	495,726	9	1	6	15
GDP/Worker	County	493,617	63,095	9,846	46,835	136,763
Population density	Municipality	495,726	282	382	7	4,653
AfD vote share (2017) (in %)	Municipality	492,618	15	7	3	45
Share high school (in %)	Municipality	495,726	29	8	0	58
Share broadband access (in %)	Municipality	495,726	83	11	44	100
Share immigrants (in %)	Municipality	483,072	14	8	2	50
Asylum Seekers/Pop.	County	495,726	0.011	0.006	0.000	0.102

Notes: This table reports summary statistics for the main variables in the estimation sample. Variables tagged with a † are scaled by population (in 1,000).

2.1 Anti-Refugee Incidents

The data on incidents targeting refugees were collected by the Amadeu Antonio Foundation and Pro Asyl (a pro asylum NGO).[3] These data cover incidents including anti-refugee graffiti, arson of refugee homes, assault, and incidents during protests in Germany between January 2015 and early 2017. This period is of particular interest since it includes the beginning and height of the refugee crisis in Germany. All 3,335 anti-refugee aggressions feature a short description and are classified into four groups. The most common cases are property damage to refugee homes (2,226 incidents), followed by assault (534), incidents during anti-refugee protests (339), arson (225). 11 events are classified as suspected cases that were still under investigation. Table A.2 in the online appendix lists examples for each class of anti-refugee activity.

All incidents are geo-coded with an exact longitude and latitude, which we use to assign them to municipalities.[4] Figure 1 shows the location of the anti-refugee incidents in our observation period for each German municipality.

The data appear to be high quality. Each entry has a clearly indicated source. Nearly half of the incidents in the dataset are reported by the federal government in response to inquiries by the left-wing party "Die Linke". Other sources include police reports and national or local media outlets. We hand-checked a random sample of 100 incidents and found their coding accurately reflected the information reported in the respective source.

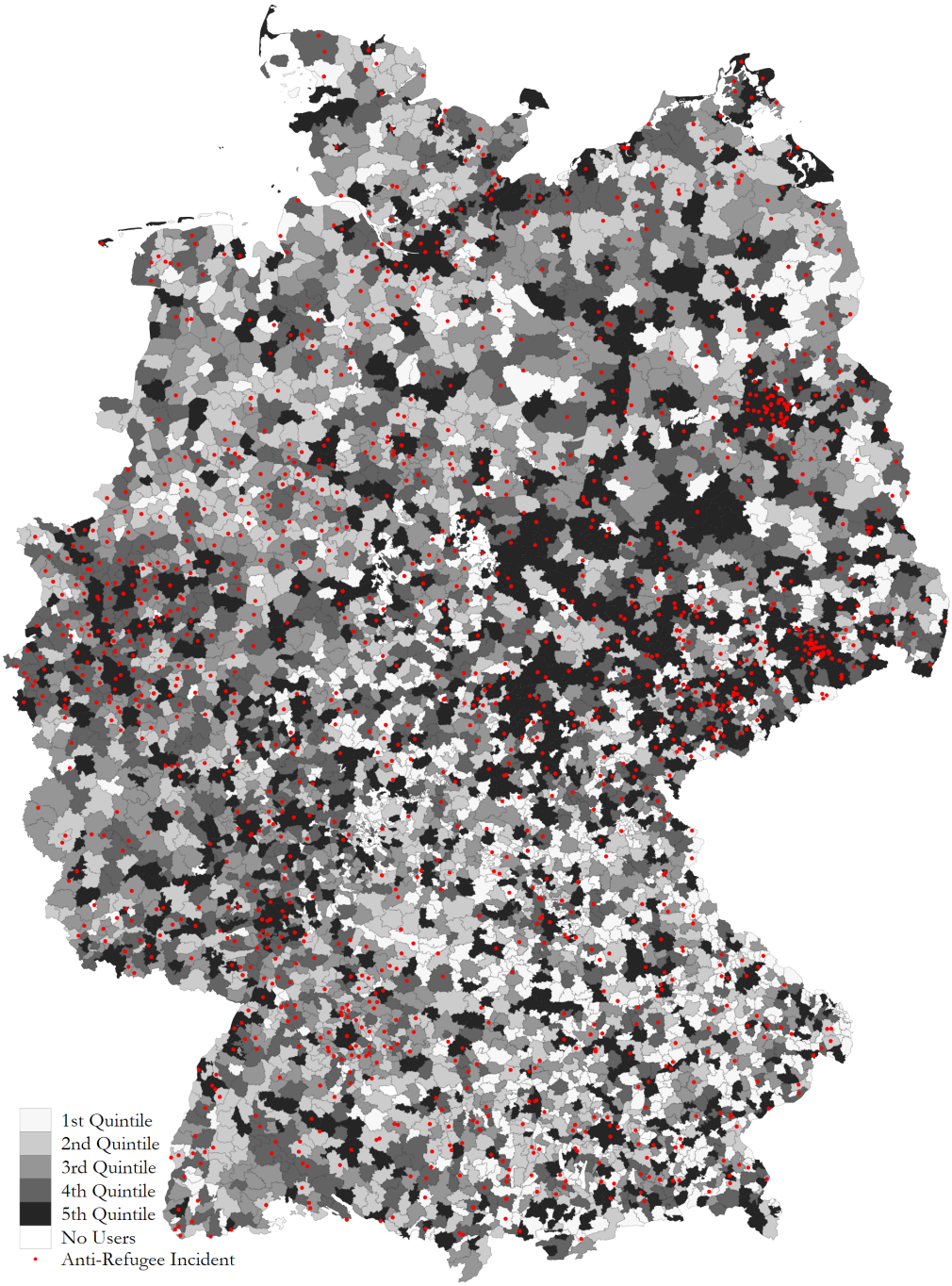
2.2 Facebook Data on Refugee Salience

We construct a proxy for the frequency of anti-refugee messaging on social media based on the Facebook page of the AfD. We chose the AfD's page because the party is by far the most popular far-right political movement in Germany. At the time of the refugee crisis, the AfD also had the highest number of Facebook followers of *any* German party. This makes their page arguably the most important platform of exchange about refugees among Germany's right-wing social media users.

[3] These data are available at `https://www.mut-gegen-rechte-gewalt.de/service/chronik-vorfaelle`.

[4] To assign coordinates to municipalities, we use the shape files provided by the ©GeoBasis-DE/BKG 2016 website. The shape file contains data for the 4,679 German municipalities ("Gemeindeverwaltungsverband"). 213 of these municipalities do not have inhabitants (e.g. forest areas) nor anti-refugee incidents. After dropping these cases, we are left with 4,466 municipalities in our estimation sample. We use the level of the "Gemeindeverwaltungsverband" since these exhibit smaller differences in their size and population than the 11,165 German "Gemeinden" and are therefore more suitable for spatial analysis according to the data provider (see link).

Figure 1: AfD Facebook Usage per Capita and Anti-Refugee Incidents



Notes: This map plots the number of Facebook users of the Alternative for Germany (AfD) page per capita for each of the 4,466 German municipalities. The red dots indicate the locations of the 3,335 anti-refugee incidents from the Amadeu Antonio Foundation.

We start by using the Facebook Graph API to collect all status posts, comments, and likes from the AfD Facebook page (see Appendix B.1. for an introduction to Facebook). The API provides a unique identifier for each post, allowing us to link posts to comments and likes, as well as the users who posted, commented, or liked anything on the page. Overall, we collected 176,153 posts, 290,854 comments, 510,268 likes, and 93,806 individual user IDs.

As our baseline measure for the salience of anti-refugee hate speech on social media, we use the number of posts on the AfD Facebook page that contain the word "Flüchtling" (refugee) in any given week. The narrative in these posts centers around the idea that the "elites"—politicians and mainstream media outlets—have betrayed "the people" by allowing "streams" of illegitimate "economic refugees" to enter the country, who are described as being criminals and rapists for "cultural reasons". Table A.1 in the online appendix provides a few representative examples; Section 3.5 provides a more in-depth analysis. A potential downside of this approach is that we may inadvertently tag posts that do not express negative sentiments towards refugees. However, a careful content analysis of posts and comments reveals that the overwhelming majority appear to agree with the positions of the AfD. This is perhaps unsurprising given that only people who "like" the AfD Facebook page will be informed about new posts. Critics, on the other hand, have a strong incentive not to indicate publicly that they "like" the party.

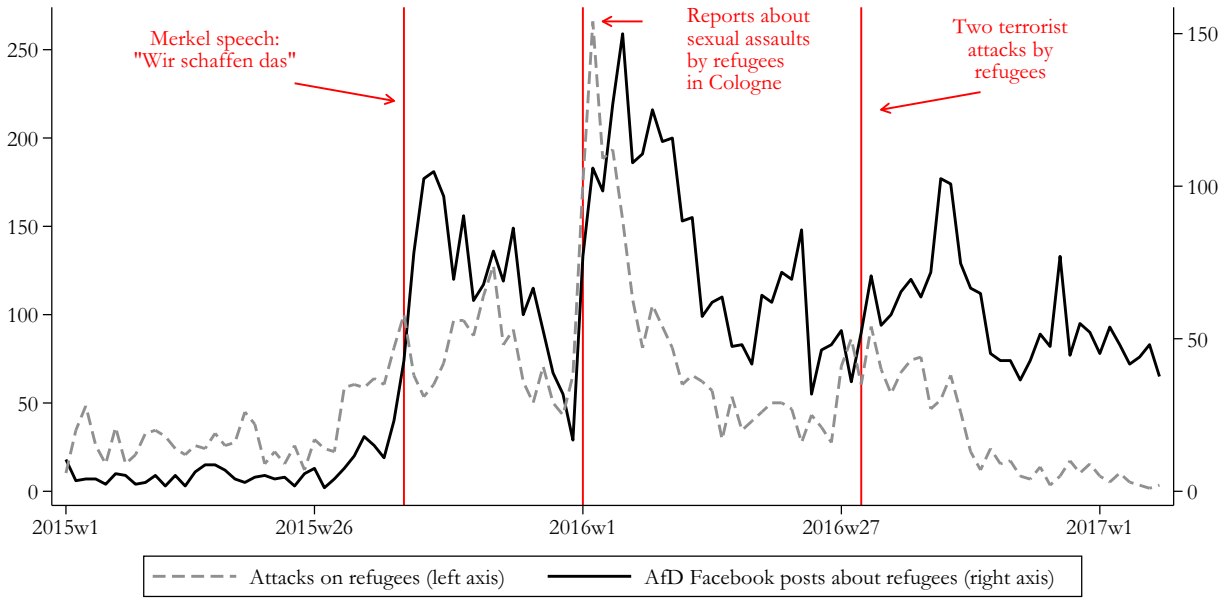
We plot the total number of AfD Facebook page posts about refugees and the number of anti-refugee incidents in Figure 2. Weeks with more refugee posts also tend to have more anti-refugee events. Both series clearly spike during salient events related to refugees, such as Angela Merkel's widely reported statement "Wir schaffen das" ("We can do this") during a press conference on the challenges of the refugee situation. A simple time series regression of refugee attacks on AfD posts yields a R^2 of 0.34 (unreported).

2.3 Municipal-Level Facebook Measures

We construct a measure of exposure to right-wing social media at the municipal level. Because survey data about German Facebook usage are, to our knowledge, only available at the level of the 16 federal states, we hand-collect user location data by using the unique user identifiers provided by the Facebook Graph API. Due to Facebook's privacy policy, we are only able to collect this information for people who make it publicly available.

Because we are interested in the transmission of right-wing social media sentiment, we measure exposure to it on Facebook based on users of the AfD page. In total, we can identify

Figure 2: Refugee Posts on Social Media and Anti-Refugee Incidents Over Time



Notes: This figure plots the number of posts about refugees on the Facebook page of the "Alternative for Germany" and the number of anti-refugee incidents in Germany over time.

93,806 users who interacted with the page at least once.[5] We were able to hand-collect and geocode a place of residence for 34,396 of these users. Overall, we were able to identify at least one AfD Facebook page user for 3,563 of the 4,466 municipalities.[6] In Figure 1 we visualize the distribution of AfD users per capita. Anti-refugee incidents are concentrated in areas with more right-wing social media users. To illustrate this, Figure A.2 in the online appendix shows the share of municipalities with at least one refugee attack, depending on whether we can identify *at least one* AfD Facebook page user. Municipalities with AfD users are three times as likely to experience an attack during our observation period. Out of the total 3,335 attacks on refugees in our sample, 3,171 occurred in municipalities with AfD Facebook page users. A t-test rejects the null hypothesis of no difference between the mean of the two groups with a value of 22.95.

Using the location data for AfD users, we can also assign posts, comments, and likes to municipalities. Based on these data, we construct auxiliary measures of social media interactions,

[5]The Facebook API does not provide data on which users "like" a page but only on users who *interact* with a page, e.g. by liking another user's comment. As a result, the total number of user IDs we have is smaller than the more than 300,000 people who had liked the AfD Facebook page as of 2017.

[6]Note that the decision of users to disclose their location is unlikely to matter in our setting. This is because we exploit variation *within* the same location over time, which abstracts from time-invariant endogenous selection using municipality fixed effects.

e.g. the number of local posts scaled over the number of AfD users.[7]

2.4 Data on Internet and Facebook Outages

We collect data on local internet outages from Heise Online. Heise lists user reports of internet problems by telephone area codes and includes start times and duration. We use area codes to assign internet problems to municipalities; the start date and duration allow us to count the number of problems for each municipality and week.[8] The internet outage reports are geographically dispersed with no clear patterns of regional clustering (see Figure A.4a). The outages are also dispersed over time Figure A.4b.

To validate the Heise data, we search for newspaper reports on major internet disruptions. While the large-scale and short-lived outages discussed in the newspaper reports are not representative of the more localized and longer-lasting outages we exploit in our regressions, they do suggest that the Heise data provide a valid proxy for internet disruptions. For all major disruptions we could identify in newspapers, the Heise data suggest an increase in the number of outages specific to the internet provider experiencing the outage. Table A.5 lists several examples of newspaper reports on such outages and the respective information in our data.[9]

We focus on major outages that fulfill two criteria: (1) they have to last longer than 24 hours, and (2) they affect a significant part of the population (be in the top quartile of the reported internet problems to population ratio). This gets around the issue that some reports may reflect individual users' glitches rather than general disruptions.[10]

We also collect information on major Facebook disruptions. To identify these, we start by searching for newspaper reports of Facebook problems in our sample period. In total, we find reports on eight large outages (see Table A.6 for an overview and more details). We then validate their precise timing using the number of weekly user-reported Facebook problems on

[7]We find that some users post and comment excessively, which leads to a few outliers in measuring how active users are in a given municipality. We therefore winsorize the number of posts, comments, and likes we can attribute to local users at the 99.9th percentile to avoid individual users driving the results.

[8]If an area code spans multiple municipalities, we assign an internet outage to the municipality that overlaps most with the area code. We prefer this over to assigning the outage to all municipalities within the area code's territory because some area codes include minor overlaps with many municipalities. Assigning an internet outage to all of these municipalities would introduce substantial noise.

[9]To interpret the number of outages, note that the Heise data reports an average of four reported internet outages per provider per week. That means even an increase of 15 reported outages represents a large increase.

[10]In some cases, users do not seem to report the end date of the internet outage, which can lead to unlikely durations of several months. We thus winsorize the maximum duration at 3 weeks, but this choice is not material for our results. We scale outages over population because towns with more inhabitants mechanically also report more disruptions. As we discuss below, our results are robust to using alternative definitions of this cut-off.

the website of "Allestörungen", a portal for aggregating user complaints on individual websites and apps. Perhaps unsurprisingly, the eight outages widely reported on in the news media are also associated with spikes in user-reported problems.

Using these data, we define a dummy variable that is 1 for weeks with Facebook outages and 0 otherwise. These outages have the advantage that they are specific to Facebook; in fact, they are uncorrelated with the total number of weekly internet outages in a given week from our Heise data. In contrast to the internet disruptions, the downside is that Facebook outages are rare, shorter, and only generate weekly variation.

2.5 Auxiliary and Control Variables

We obtain control variables from a host of sources, which are explained in more detail in the online appendix. Socioeconomic data on the municipality and county level are from the German Statistical Office, available via www.regionalstatistik.de. We include information on each municipality's population by age group, GDP per worker, population density, the share of the population with a high school degree ("Abitur"), the share of the population receiving social benefits, the share working in manufacturing, and the vote results for the 2017 German Federal Election. To control for "pull factors" of anti-minority crimes, we also obtain the share of the population that are immigrants and asylum seekers.

To measure the extent to which people use the internet, we use the share of households in a county with broadband access as well as average mobile download speeds, collected by the Federal Ministry of Transport and Digital Infrastructure (BMVI).[11] In addition, we use the number of registered *.de* internet domains per capita in a county to measure internet affinity, which has a correlation of 0.48 with broadband access.

To measure the local penetration of traditional media, we obtain data for 2016/2017 newspaper sales from the "Zeitungsmarktforschung Gesellschaft der deutschen Zeitungen (ZMG)" (Society for Market Research of German Newspapers).[12] Based on this data, we construct a measure of traditional newspaper consumption as the number of newspaper sales per capita.

[11] Broadband access is highly correlated with publicly available survey data on individuals' internet use from Eurostat; these data are only available on the state level (see Figure A.3 in the online appendix).

[12] These data contain the number of print newspapers sold in each municipality with more than 3,000 inhabitants. Newspapers are listed if, in any given town, they (1) sell at least 50 copies and (2) have a market share of at least 1%. To have a similar sample size across specifications, we impute values for 1,120 towns for which news paper sales data are not available, based on a municipality's population, population density, AfD vote share, and county fixed effects. However, the results are almost equivalent without imputation (available upon request).

For our comparison of social and more traditional media, we collected the number of total and refugee-related reports in German news media from Nexis UNI (previously LexisNexis). We use this to construct the weekly share of news reports about refugees. For further analysis, we obtained the full text of all refugee-related reports using the Lexis bulk data API, as well as all Facebook data from the pages of five major German newspapers (Welt, Frankfurter Allgemeine Zeitung (FAZ), Tageszeitung (TAZ), Süddeutsche Zeitung (SZ), and Bild).

We also include controls for the local prevalence of right-wing extremism. One such measure is the number of murders committed by neo-Nazis in each municipality from 1990 until 2016, which were collected by "Mut gegen rechte Gewalt" (Courage Against Right-Wing Violence). We complement this proxy for contemporary right-wing violence with data on the historic prevalence of anti-semitism collected by Voigtlander and Voth (2012).[13]

Finally, we obtain Google trends data on overall interest in the search terms "Brexit", "Trump", and "UEFA EM 2016" in Germany to proxy for distracting news events. Google scales the weekly number of searches for these terms on a scale from 0 to 100, where 100 marks the week with the highest search interest in the preceding 5 years. The time series plots in Figure A.8 in the online appendix suggest these measures are sound approximations for attention paid to Brexit, the Trump election, and the UEFA European Championship (one of the most widely followed sports events in Germany).

3 Empirical Strategy and Main Results

3.1 Empirical Strategy

We begin to investigate the link between social media and anti-refugee incidents by estimating fixed effects panel regressions akin to a Bartik-type approach (Bartik, 1991). In particular, we use the interaction of local right-wing Facebook usage ($AfD\ Users/Pop_i$) and weekly refugee posts on the AfD Facebook page ($Refugee\ Posts_t$) to measure the differential change of hate crimes conditional on anti-refugee sentiment on social media. This empirical set-up creates variation by week and municipality, which we exploit in the following regression model:

[13] From their dataset, we use the natural logarithm of one plus the number of deported Jews as well as one plus the number of letters written to "Der Stürmer", the antisemitic newspaper published by Nazi politician Julius Streicher. Towns with no information are coded as zero. We do not use scaled variables because the data from Voigtlander and Voth (2012) only cover a fraction of the municipalities in our sample.

$$
\begin{aligned}
Refugee\ attack_{it} = {} & \beta\ AfD\ Users/Pop_i \times Refugee\ Posts_t \\
& + \gamma\ Controls_i \times Refugee\ Posts_t \\
& + Week\ FE_t + Municipality\ FE_i + \epsilon_{it},
\end{aligned} \tag{1}
$$

The dependent variable is a dummy for the incidence of a refugee attack in municipality i in week t. β measures the differential change in anti-refugee incidents conditional on Germany-wide posts about refugees on the AfD page—as a proxy of Germany-wide anti-refugee sentiment on social media—and right-wing social media users per capita. We control for a host of local characteristics interacted with the refugee post measure. Because we include many fixed effects and interaction terms, we estimate 1 using Ordinary Least Squares, which yields the linear probability model. Standard errors are clustered by municipality. We consider alternative specifications of the dependent variable and standard errors in robustness exercises.

This framework has three key features. First, it circumvents reverse causality, because refugee incidents in one town are unlikely to change anti-refugee sentiment in *all other* towns. Second, our measure of social media exposure is time-invariant and thus not the result of whether a municipality experiences refugee attacks in a given week.[14] Third, a full set of fixed effects controls for unobserved heterogeneity that affects all towns at the same time (such as salient news events), as well as time-invariant differences across towns (such as a history of anti-minority violence).

The main concern with estimating Equation (1) is that $AfD\ Users/Pop.$ may be correlated with other municipality characteristics that could explain differences in how local anti-refugee attacks co-vary with the salience of refugees online. In that case, we would not be capturing a pure social media "effect". For example, the share of AfD Facebook subscribers may partially pick up general right-wing attitudes, which could lead to more anti-refugee attacks in times of high refugee salience. This concern may also not be sufficiently addressed by controlling for interactions of observable municipality characteristics with the refugee salience measure.

We therefore develop an identification strategy based on Facebook and internet outages. These disruptions induce plausibly exogenous variation in people's exposure to social media while leaving other local characteristics unchanged. The first part of this empirical strategy exploits the timing of major server problems at Facebook, which disrupt access to the platform.

[14] In the robustness section below, we alternatively measure local social media penetration before the start of the refugee crisis, at the cost of reducing the number of users for whom we have location data. This adjustment makes little difference for the results.

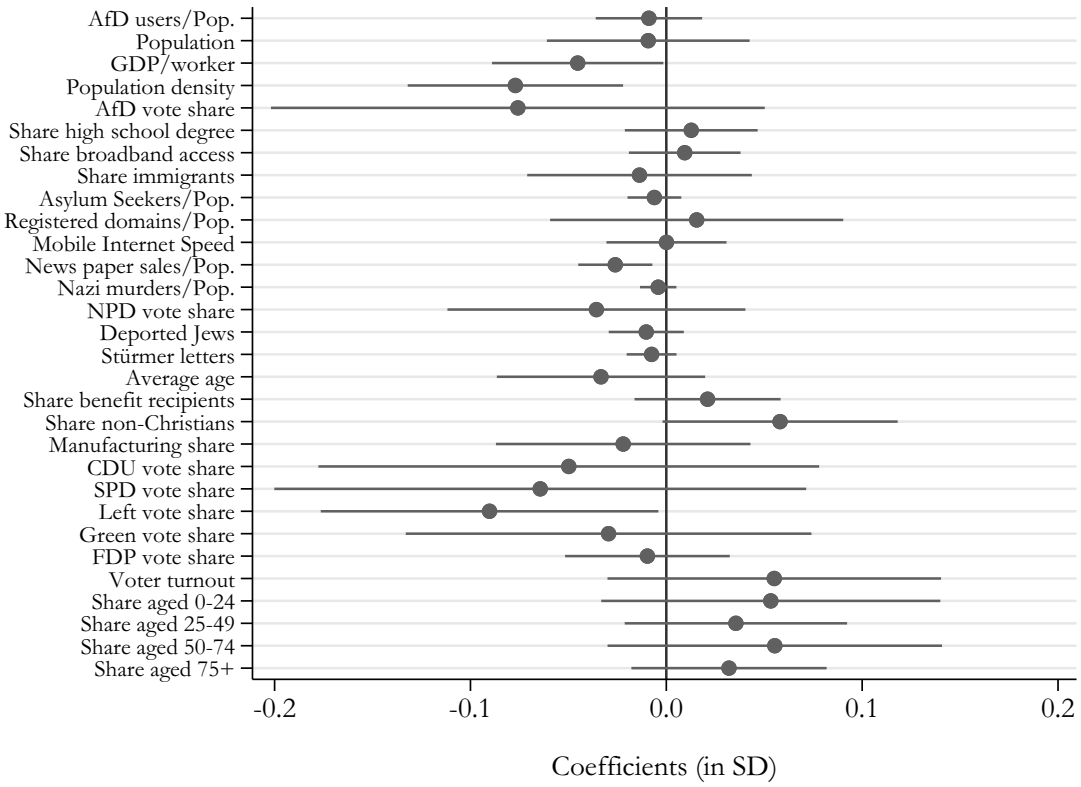
In the second part, we build on the insight that German internet infrastructure is trailing behind that of many other European Countries (e.g. Latvia) and the OECD average (see Financial Times, 2017; OECD, 2016). As a result, prolonged internet outages are relatively common. Because around 50% of worldwide Facebook users accessed the platform with their computers, many users are exposed to disruptions in internet access. In Germany, this share is likely to be even higher because of the relatively slow adoption of mobile internet.[15]

Local internet outages are widely dispersed geographically: Figure A.4a visualizes the distribution of disruptions per capita across Germany. The outages are also not particularly clustered in a particular time period (see Figure A.4b). Crucially, the frequency of internet problems is uncorrelated with the share of the population on the AfD Facebook page. As such, internet disruptions provide exogenous variation that is not already captured by our variable on local Facebook usage. The number of reported internet problems is also uncorrelated with the total number of refugee attacks in a given municipality. In fact, regressing the frequency of internet outages on a host of municipality characteristics in Figure 3 suggests that they are largely uncorrelated with observable factors: the estimated coefficients are nearly all statistically indistinguishable from zero and quantitatively small. Taken together, our interpretation is that whether an internet outage occurs in a given town and week is as good as randomly assigned with regard to unobserved other factors that might drive hate crimes.

We analyze the effect of Facebook and internet outages in a flexible empirical framework. We begin by asking whether these outages reduce anti-refugee attacks, and whether they do so particularly in areas with a higher concentration of AfD Facebook users. We then study whether these disruptions also decrease our baseline correlation of local exposure to anti-refugee sentiment and hate crimes. More formally, the most saturated regressions have the following triple difference form:

[15]Data on Facebook usage patterns reported on Statista.com and on mobile internet usage in Germany on (also on Statista.com) support this assessment.

Figure 3: Balancedness — Internet Outages and Local Characteristics



Notes: This figure plots the coefficients of the regression $\overline{Internet\ outages}_i = \alpha + \mathbf{X}'\beta + \epsilon_i$, where the dependent variable is the total number of internet outages in a municipality (based on our baseline definition) and $\mathbf{X}$ is a vector of local characteristics for which we plot the estimates. To make the magnitudes comparable, we standardize all variables to have a mean of zero and standard deviation of one. 95% confidence intervals are based on standard errors clustered by municipality.

$$
\begin{aligned}
Refugee\ Attack_{it} = {} & \beta\ AfD\ Users/Pop_i \times Refugee\ Posts_t \\
& + \lambda\ Outage_{it} \times AfD\ Users/Pop_i \times Refugee\ Posts_t \\
& + \delta_1\ Outage_{it} + \delta_2\ Outage_{it} \times Refugee\ Posts_t \\
& + \delta_3\ Outage_{it} \times AfD\ Users/Pop_i \\
& + \gamma_1\ Controls_i \times Refugee\ Posts_t \\
& + \gamma_2\ Controls_i \times Outage_{it} \\
& + Week\ FE_t + Municipality\ FE_i + \epsilon_{it},
\end{aligned} \tag{2}
$$

For the Facebook outages, which only vary by week, we replace $Outage_{it}$ with $Outage_t$.[16] For the initial tests, we focus on the estimates for δ_1 and δ_3 while excluding the coefficients β, λ, δ_2, and γ_1. That is, we ask whether outages reduce anti-refugee incidents, and whether they reduce them more in areas with more AfD Facebook users. In the fully interacted regressions, the main coefficient of interest λ captures the correlation of anti-refugee attacks and local exposure to anti-refugee sentiment on social media, depending on whether an outage occurs. Put differently, we test whether outages break the correlation between real-life incidents and refugee salience, particularly for areas with high right-wing Facebook penetration. The vector $Controls_i \times Outage_{it}$ controls for the differential effect of outages based on observable characteristics, such as internet affinity.

The identifying assumption of this approach is that Facebook and internet outages only affect anti-refugee incidents through their effect on social media exposure. This assumption is plausible for Facebook outages. In the case of internet outages, for which we have variation at the municipality-week level, one may be worried about alternative online channels. We discuss these and other potential threats to identification in the next section.

Exploiting variation in Facebook and internet outages also allow us to address the concern that towns with a stronger right-wing presence may show differential trends whenever the nationwide sentiment towards refugees changes. This is because these relatively short-lived outages are unlikely to affect the presence of deep-rooted right-wing attitudes in a municipality; absent online channels, the outages should thus not have an impact on real-life outcomes. The framework in Equation (2) further addresses reverse causality concerns. If we were merely capturing that local incidents drive posts on social media, Facebook and internet outages should

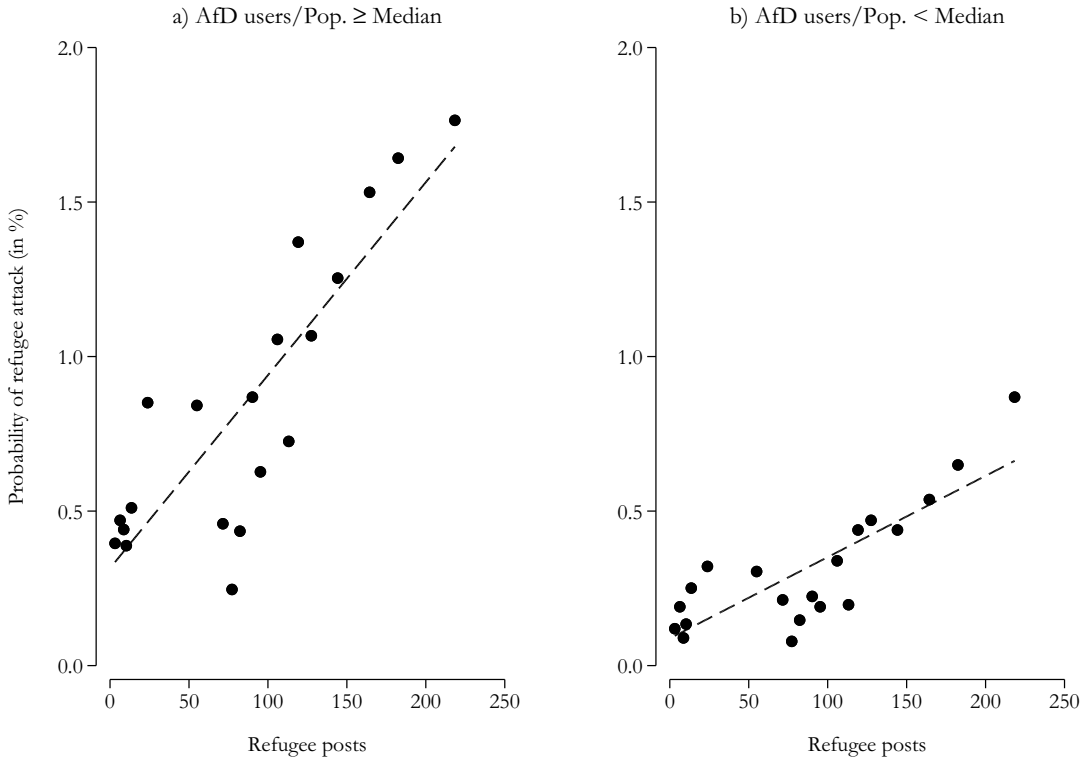
[16]Note that, as a result, the estimates of δ_1 and δ_2 in Equation (2) are absorbed by the week fixed effects.

not reduce the number of hate crimes. Instead, they should only reduce social media activity, keeping the number of anti-refugee incidents unchanged.

3.2 Panel Regression Results

We illustrate the intuition behind our regression framework in Figure 4. The figure shows a binned scatter plot of anti-refugee attacks and anti-refugee sentiment, split by the degree of exposure to right-wing social media. Higher refugee salience is associated with a higher probability of anti-refugee attacks in both sub-samples, but the positive slope is far more pronounced for towns with an above median AfD user to population ratio (Panel (a)). Our baseline regression coefficient picks up the difference in slopes between municipalities with high and low Facebook usage.

Figure 4: Exposure to Refugee Sentiment on Facebook and Hate Crimes



Notes: This figure plots the average number of anti-refugee attacks against our measure of anti-refugee sentiment for municipalities below and above the median of $AfD\ Users/Pop.$ Refugee attacks are binned by 20 quantiles of refugee posts and residualized with respect to population.

Table 2 presents the regression results from estimating Equation (1) with varying sets of control variables (interacted with refugee salience). The coefficient on the interaction of

local Facebook usage and Germany-wide refugee posts is positive and highly significant in all specifications. Column 1 shows the panel regressions with the baseline control variables, which yields a coefficient 0.024 on the interaction term. This correlation does not appear to be driven by support for the AfD alone: the result holds although we control for the AfD vote share in the 2017 federal election. This highlights a distinction between our social media measure and general support for the party.

To get a sense of the magnitudes, consider as a case study the cities of Bochum and Hannover, which are about one standard deviation apart in the ratio of AfD users to population (in 1000s) (≈ 0.29). Holding average anti-refugee sentiment in our data constant (84 posts), this means a one standard deviation higher right-wing social media usage is associated with a 10% higher probability of an anti-refugee incident relative to the mean. Table A.12 in the online appendix shows that this correlation is largely driven by cases of assault.

In columns 2 through 6, we introduce a richer set of controls that accounts for local right-wing attitudes, general media exposure, more socio-economic factors, and the vote shares of all major parties in the 2017 election (see Table A.3 for an overview of the control variables). In column 7, we add all interacted controls jointly. The inclusion of these covariates makes little difference to our main estimate. This is a first indication that the correlation between social media exposure and anti-refugee incidents is not driven by observable municipality differences unrelated to Facebook usage.

3.3 Quasi-Experimental Evidence: Facebook and Internet Outages

To isolate the importance of social media, we next draw on internet and Facebook outages as sources of quasi-experimental variation. To count as a severe internet disruption, our baseline measure has to fulfill two criteria: (1) it has to last at least 24 hours, and (2) it has to affect a significant part of the population, i.e. be in the top quartile of reported internet disruptions per capita, which vary by municipality and week (see section Section 2 for more details). This gives us 313 severe internet outages.[17]

Internet outages. Are local internet outages severe enough to decrease a municipality's exposure to social media? We investigate this question by using a sample of posts from the AfD

[17] In the online appendix, we show our results are robust to alternative definitions. We also exploit the eight major Facebook outages, which only vary by week. We discuss the results and their interpretation in turn.

Table 2: Baseline Correlations — Facebook Posts and Hate Crime

		Additional interacted controls					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Baseline controls	Right Wing controls	Media controls	Socio-economic controls	2017 vote controls	Age structure controls	All controls
AfD users/Pop. × Refugee posts	0.024***	0.020**	0.023**	0.024**	0.021**	0.023**	0.016**
	(0.009)	(0.008)	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)
Observations	479,964	479,964	479,964	474,303	479,964	476,856	474,303
R-squared	0.082	0.083	0.082	0.083	0.083	0.083	0.084
Municipalities	4324	4324	4324	4273	4324	4296	4273
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline controls [8] × Posts	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Right-wing controls [4] × Posts		Yes					Yes
Media controls [4] × Posts			Yes				Yes
Socio-econ. controls [4] × Posts				Yes			Yes
Election controls [7] × Posts					Yes		Yes
Age controls [4] × Posts						Yes	Yes

Notes: This table presents the estimated coefficients from a regression of hate crimes against refugees on the interaction of local social media usage and anti-refugee sentiment as in Equation (1). The dependent variable is a dummy for the incidence of a refugee attack. *AfD users/Pop.* is the ratio of people with any activity on the AfD Facebook page to population. *Refugee posts* is the Germany-wide number of posts on the AfD's Facebook wall containing the word refugee ("Flüchtling"). All control variables are interacted with the *Refugee posts* measure; see text for a description of the controls. Robust standard errors in all specifications are clustered by municipality. ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.1 levels, respectively.

Facebook page for which we know the users' locations.[18] Figure 5a plots the local number of posts against the intensity of local internet outages. Local Facebook activity falls with outage intensity and is close to 0 as soon as we observe more than 0.25 outage reports per 10,000 inhabitants. Figure A.5 shows that we observe significantly fewer posts and comments on Facebook for municipalities that experience an internet disruption. These results lend credence to the idea that exposure to social media content is reduced in the affected municipalities and not compensated by users accessing Facebook with their mobile phones.

If internet outages indeed reduce local social media exposure, we would expect them to mediate the capacity of social media to propagate anti-refugee incidents. As described in Section 3.1, we test this hypothesis by interacting the main terms of interest $AfD\ Users/Pop_i \times Refugee\ Posts_t$ with $Internet\ Problems_{it}$, our dummy for severe internet disruptions. We graphically illustrate the results in Figure 5b. The binned scatter plot is almost identical to Figure 4, except that we plot a separate slope for municipalities that experience an internet outage. This reveals a striking pattern: while anti-refugee attacks increase with anti-refugee posts, this relationship disappears in municipalities that experience an internet outage. This holds true for municipalities with high and low Facebook usage.

Figure 5b implies that internet outages have a substantial attenuating effect. Consider the pattern in panel (a). Without outages, there is a strong correlation of refugee posts and attacks. During outages, the correlation is essentially zero. This means that the outage effect is larger than the baseline estimate for $AfD\ Users/Pop. \times Refugee\ posts$, which is given by the slope difference of the dotted lines in panels (a) and (b). We interpret this as evidence that cutting of users from social media completely has large effects.

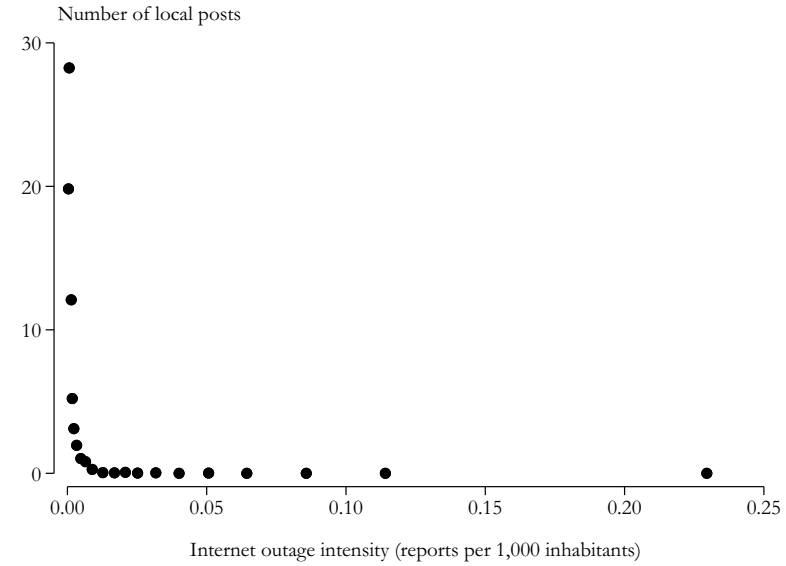
We next estimate versions of Equation (2) and report the regression results in Table 3. Column 1 shows that internet outages reduce anti-refugee violence. The coefficient of -0.003 implies that, during such outages, the probability of a refugee attack is 53% lower relative to the dependent variable mean (≈ 0.006). In Figure 6, we investigate the timing of this drop in incidents. Because the outages are relatively rare in the municipality-week panel, the estimates are necessarily noisy. Nonetheless, we can see a reduction in anti-refugee incidents that is sharply concentrated in the week of the internet outage.

Column 2 in Table 3 implies that this effect is driven by periods of high sentiment; it may also be driven by areas with many AfD Facebook users (column 3) but the coefficient is not

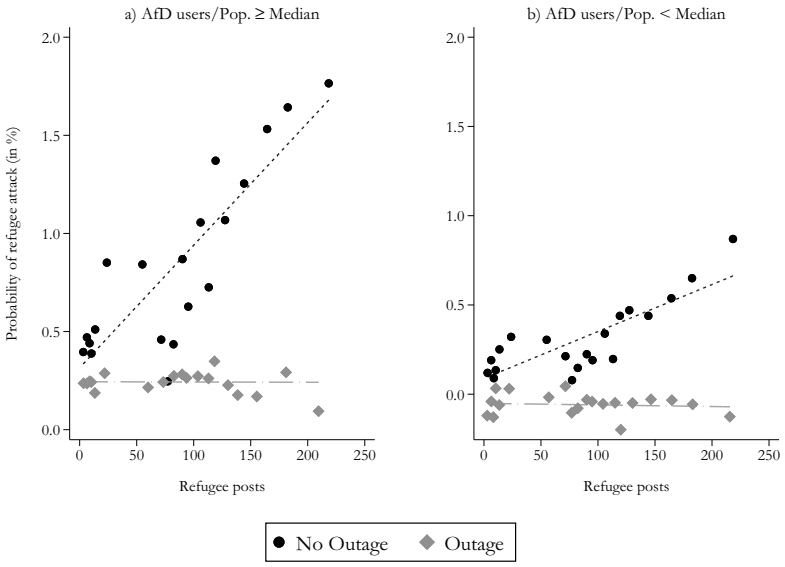
[18] These posts and comments are a sub-sample by users who publicly disclosed their location in their Facebook profiles.

Figure 5: Quasi-Experimental Results from Internet Outages

(a) Internet Outages Reduce Local Facebook Activity

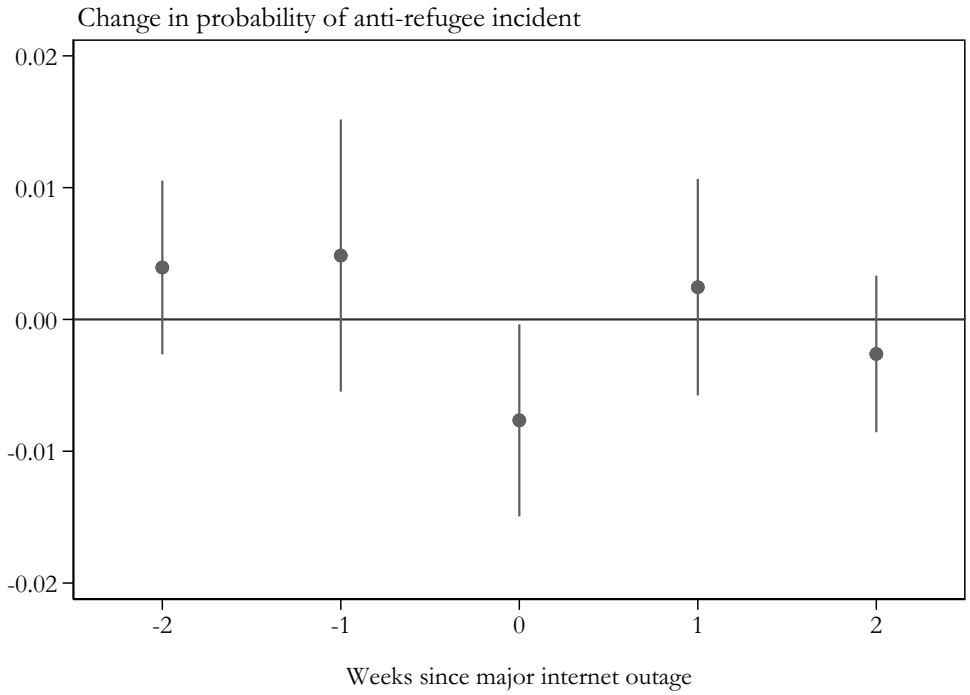


(b) Internet Outages Reduce Local Anti-Refugee Incidents



Notes: Panel (a) shows a binned scatter plot of local posts on the AfD Facebook page as a function of the reports on internet outages in a given week. Panel (b) plots the average number of anti-refugee attacks against our measure of anti-refugee sentiment for municipalities above and below the median of $AfD\ Users/Pop$. Refugee attacks are binned by 20 quantiles of refugee posts. We additionally split towns by whether they experience an internet outage in a given week (gray squares). The number of anti-refugee attacks is residualized with respect to population; hence, the number of attacks can be slightly below 0 in some bins.

Figure 6: Internet Outage Event Study



Notes: This figure plots estimates the estimates for δ from the event study regression $Attacks_{it} = \sum_{t=-2}^{2} \delta_{w=t} Outage_{it} + Fixed\ Effects + \epsilon_{it}$, where $Outage$ refers to internet outages in municipality i in week t. 95% confidence intervals are based on standard errors clustered by municipality.

statistically significant. In columns 4 through 6, we estimate the full triple-difference model. Here, we estimate the effect of outages in areas with high social media use at times of high anti-refugee sentiment. The estimates suggest that internet problems reduce social media's impact on anti-refugee violence. While the coefficient of refugee posts and social media exposure is similar to our baseline correlations, the triple interaction term with internet outages is negative and statistically significant in all three specifications. Quantitatively, internet outages appear to mitigate the entire effect of social media. In line with the graphical evidence in Figure 5b, we find that the triple interaction coefficient is larger than the baseline coefficient. Put differently, for a given level of anti-refugee sentiment, there are fewer attacks in municipalities with high Facebook usage during an internet outage than in municipalities with low Facebook usage *without* an outage.

Could it be that the effect of internet outages is merely coincidental? As an alternative way of assessing statistical significance, we perform a randomization test. Instead of the actual internet disruptions, we randomly define 313 municipality-week pairs as placebo outages. We then estimate the same regression using 500 different sets of placebo outages. This allows us to evaluate the probability of finding a statistically significant coefficient in our dataset. Using this procedure, we find that more than 99% of the placebo triple interaction coefficients exhibit a lower t-statistic than our estimate. Our findings are thus unlikely to be purely coincidental. We show the full distribution of t-statistics from this randomization test in Figure A.7a in the online appendix.

The identifying assumption for internet outages in our framework is that they only have an effect on anti-refugee hate crime through the reduced exposure to social media. Could it be that we observe reduced hate crimes because users are cut off from the internet generally, not from social media in particular? Two pieces of evidence support the idea that we capture a social media channel.

First, when we include interactions of internet disruptions with measures of internet usage (broadband access, per capita internet domains, mobile internet access), our main coefficient is unaffected (see column 6 in Table 3). The coefficients of the internet usage interactions are generally statistically insignificant or have the opposite of the expected sign. This is at least some indication that we are not merely capturing general internet usage. It also suggests that our findings are unlikely to capture that people are busy fixing internet access problems. If we were merely capturing such displacement effects, one would expect it to more strongly affect people in areas with high internet usage, which does not seem to be the case in the data. Second,

Table 3: Local Internet Outages and Social Media Transmission

	(1)	(2)	(3)	(4)	(5)	(6)
Baseline Interaction						
AfD users/Pop. × Refugee posts				0.024***	0.016**	0.016**
				(0.009)	(0.008)	(0.008)
AfD users/Pop. × Posts × Outage				-0.181***	-0.184***	-0.172***
				(0.058)	(0.058)	(0.057)
Outage Interaction						
Outage	-0.003***	-0.000	-0.003**	-0.001	-0.002	-0.007
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.008)
Refugee posts × Outage		-0.005***		-0.000	0.001	0.000
		(0.001)		(0.002)	(0.002)	(0.002)
AfD users/Pop. × Outage			-2.685	4.441	4.455	4.391
			(3.464)	(4.384)	(4.054)	(4.058)
Internet Usage Interaction						
Share broadband access × Outage						-0.000
						(0.000)
Internet domains/Pop. × Outage						0.021*
						(0.012)
Mobile Broadband Speed × Outage						0.000
						(0.000)
Observations	479,964	479,964	479,964	479,964	474,303	474,303
R-squared	0.082	0.082	0.082	0.082	0.084	0.084
Municipalities	4324	4324	4324	4324	4273	4273
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Baseline controls [8] × Posts	Yes	Yes	Yes	Yes	Yes	Yes
All other controls [22] × Posts					Yes	Yes

Notes: This table presents the estimated coefficients from a regression of hate crimes against refugees on the interaction of local social media usage and anti-refugee sentiment as in Equation (1). The dependent variable is a dummy for the incidence of a refugee attack. *AfD users/Pop.* is the ratio of people with any activity on the AfD Facebook page to population. *Refugee posts* is the Germany-wide number of posts on the AfD's Facebook wall containing the word refugee ("Flüchtling"). Internet outages are defined as municipality-weeks that are in the top quartile of the ratio of reported internet outages to population. The coefficient of "Refugee posts × Outage" is multiplied by 100 for readability. Columns 1-4 include the baseline controls. Columns 5 and 6 include all controls as in column 7 of table 2, interacted with *Refugee posts* (unreported). Column 6 further adds the interaction of broadband access and internet domains/pop. with local internet outages. Robust standard errors in all specifications are clustered by municipality. ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.1 levels, respectively.

after including the other interaction terms in columns 4 through 6, the coefficient on internet outages is no longer statistically significant. This result also supports the idea that internet outages reduce hate crime by limiting access to social media.

Another concern could be that hate crimes are less likely to be reported during internet outages. We believe this is unlikely to explain our findings because we analyze incidents that happened years in the past. While internet outages might hamper the flow of information, it seems highly unlikely that incidents such as assault or property damages are *never* reported due to a temporary internet disruption. As further evidence, we limit our analysis to official reports by the police or the German parliament, for which social media reporting is an unlikely concern. This yields similar results (see column 1 of Table A.8).

We also run a number of tests to rule out that our Germany-wide measure of refugee posts is affected by local internet outages. As stated above, this appears unlikely because we focus on *local* disruptions to the internet; Table A.7 in the online appendix shows that the total number of internet outages in a given week is uncorrelated with the total number of Facebook posts. The outage results are also robust to using a leave-one-out measure of refugee posts (column 2), Germany-wide posts in the previous week (column 3), and an alternative measure based on Google search intensity for the word refugee (*Flüchtling*) in column 4. The implied magnitudes are almost equivalent.[19] This suggests that the outage effect is driven by exposure rather than the production of anti-refugee content. In Table A.10, we show additional robustness checks for alternative transformations of the dependent variable. The findings remain robust throughout. Table A.11 shows that the results also hold using alternative definitions of the outage dummy.

Facebook outages. As further evidence for the social media transmission mechanism, we use eight major Germany-wide Facebook outages as a source of exogenous variation specific to social media access. Table A.6 outlines the details of each of the eight outages and links to relevant press reports. By definition, these outages are Facebook-specific and therefore do not affect other potential channels of online transmission.

Table A.7 in the online appendix shows that these outages are large enough to disrupt weekly activity on right-wing social media. Column 1 and 2 show that, during weeks with Facebook outages, there are on average 11% fewer new total posts and 24% fewer posts about

[19] To see this, consider the effect implied by dividing the triple interaction coefficients by the standard deviation of these salience metrics. This suggests that internet outages have a mediating effect of 9.6, 10.5, and 11.0 for the AfD posts about refugees, the leave-one-out measure, and Google trends, respectively.

refugees on the AfD page.[20] There is no evidence of such an effect in the week before. Column 5 shows that Facebook outages are also uncorrelated with the total number of weekly internet disruptions ($t = -0.41$).

We next present the results of interacting Facebook disruptions analogous to the internet outages in Table 4. The results again reveal a clear pattern. The coefficient of -0.001 in column 1 shows that the probability of an anti-refugee incident is around 18% lower in weeks with major Facebook outages (relative to the unconditional probability of an attack). Figure A.6 suggests that the timing of this effect is concentrated in the week of the Facebook outage, without significant effects in the week before or after the outage. Because we solely rely on the weekly variation from the few major Facebook outages, the estimates are noisier than those for internet outages. Column 2 shows that, intuitively, this effect is also larger in areas with many users on the AfD Facebook page. The coefficient of 2.222 suggests that Facebook outages reduce the probability of a hate crime by 12% more for a one standard deviation increase in $AfD\ users\ /Pop.$.[21] This is additional evidence that social media *per se* might affect hate crimes.

Next, we introduce the triple interaction of Facebook outages with social media usage and our refugee salience measure. The triple interaction is negative and statistically significant in all three specifications in columns 3 through 5. Quantitatively, we find that Facebook disruptions fully undo the baseline correlation of refugee attacks and exposure to social media sentiment. For example, consider that the coefficient of $AfD\ users/Pop.$ and $Refugee\ Posts$ is 0.027 in column 4 but -0.04 on the triple interaction. This implies that, in weeks of major Facebook outages, heightened refugee sentiment is not associated with a differential increase of anti-refugee attacks in municipalities with higher Facebook usage.

It is worth noting that we would expect the Facebook outage coefficients to differ in magnitude from the internet outage coefficients. This is because Facebook outages eliminate the differential exposure *between* areas with high and low social media usage to anti-refugee posts. In contrast, internet outages further exploit variation *within* municipalities. Because within-municipality variation induced by internet outages appears to matter more in our setting, we find smaller coefficients for Facebook outages.

[20] The average number of refugee posts in the time series is around 84. The coefficient estimate of 19.880 implies an effect of Facebook outages on posts of $-19.880/84 \approx 0.24$ relative to the mean.

[21] In unreported results, we also find that the interaction of Facebook outages with refugee posts has a statistically significant negative coefficient.

We again perform a randomization test to assess the statistical significance of the Facebook outage results. We randomly assign placebo Facebook outages to eight weeks in our data, excluding the weeks in which we identified Facebook outages. We then estimate the same regression using 500 different sets of placebo outages. Using this procedure, we find that 92% of the placebo triple interaction coefficients exhibit smaller t-statistics. We show the full distribution of t-statistics from this randomization test in Figure A.7b in the online appendix. This confirms that our findings are unlikely to be a matter of coincidence.

Taken together, the evidence here suggests that the relationship of anti-refugee sentiments online and hate crimes is attenuated by Facebook and internet outages. These results are most consistent with a causal propagation effect of social media.

Table 4: Facebook Outages and Social Media Transmission

	(1)	(2)	(3)	(4)	(5)	(6)
Baseline Interaction						
AfD users/Pop. × Refugee posts			0.027***	0.027***	0.021**	0.021**
			(0.010)	(0.010)	(0.009)	(0.009)
AfD users/Pop. × Posts × Outage			-0.040*	-0.040*	-0.046**	-0.046**
			(0.021)	(0.021)	(0.022)	(0.022)
Additional Outage Coeffcients						
Outage	-0.001***					
	(0.000)					
AfD users/Pop. × Outage		-2.222*	1.164	1.164	1.367	3.230
		(1.273)	(1.833)	(1.833)	(1.862)	(1.969)
Observations	479,964	479,964	479,964	479,964	474,303	474,303
R-squared	0.079	0.082	0.082	0.082	0.084	0.084
Municipalities	4324	4324	4324	4324	4273	4273
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes
Week FE		Yes	Yes	Yes	Yes	Yes
Baseline controls [8] × Posts	Yes	Yes	Yes	Yes	Yes	Yes
All other controls [22] × Posts					Yes	Yes
All controls [30] × Outages						Yes

Notes: This table presents the estimated coefficients from a regression of hate crimes against refugees on the interaction of local social media usage and anti-refugee sentiment as in Equation (1). The dependent variable is a dummy for the incidence of a refugee attack. *AfD users/Pop.* is the ratio of people with any activity on the AfD Facebook page to population. *Refugee posts* is the Germany-wide number of posts on the AfD's Facebook wall containing the word refugee ("Flüchtling"). Facebook outages refer to weeks in which Facebook experienced considerable disruptions; see the online appendix for more details on how these are defined. Note that the other interaction terms *Outage*, *Refugee posts* and *Outage* × *Refugee posts* are absorbed by the week fixed effects in columns 3-5. Columns 1-3 include the baseline controls. Columns 4 and 5 include all controls as in column 7 of table 2, interacted with *Refugee posts*. Column 5 adds the interaction of these control variables with Facebook outages. Robust standard errors in all specifications are clustered by municipality. ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.1 levels, respectively.

In the online appendix, we conduct additional robustness exercises for our outage results. In Table A.9, we show a range of different standard errors. We also assess our results' robustness to different transformations of the refugee attack variable and estimation methods in Table A.10. Our results are similar when we use the number of attacks, log(1+refugee attacks) or the ratio of refugee attacks to asylum seekers as dependent variable. In all cases, the estimated coefficients are highly statistically significant.

3.4 Additional Results

Other Posts on the AfD Facebook Page: If the channel we uncover is indeed specific to refugees, we would expect a weaker correlation between refugee attacks and posts about other topics on the AfD Facebook page. We test this hypothesis in Table A.13, where we plot the baseline estimation with refugee posts in column 1 for convenience. We also report coefficients for standardized post measures (with a mean of zero and standard deviation of one) in square brackets to compare coefficient sizes across the different posts. Next, we estimate Equation (1) using all posts except those containing the word *refugee* ("Flüchtling") in column 2. The estimate is statistically indistinguishable from zero. We also repeat our baseline test using posts containing the words "Muslim", "Islam", or "EU"—the latter is motivated by the AfD's long-standing criticism of the European Union. For all these terms, we find no significant relationship between the number of posts and the number of attacks; all estimated coefficients are considerably smaller in standardized terms compared to the baseline measure. This shows the specificity of our refugee measure: the correlation we capture does not appear to be an artifact of general anti-minority sentiment, but rather a predictable result of increased animosities towards refugees on social media in particular weeks.

Intensive Margin of Facebook Usage: If social media works as the propagating mechanism for hate speech, we would also expect its effect to increase with how frequently users interact with the AfD Facebook page. We explore this issue empirically in Table A.14, where we interact our main interaction term with the total number of local posts on the AfD wall and the number of comments and likes on AfD posts, all scaled over the number of AfD users in a municipality.[22] These measures of usage intensity are not systematically correlated with local

[22]Note that we can only construct these measures on the intensive margin of municipalities where we can identify at least one AfD user. Our baseline results also hold in this sub-sample, which we show in Table A.19 in the online appendix.

Facebook penetration, city size, or population density. As such, they create additional variation in social media engagement across towns.

The results suggest that local engagement on Facebook matters: all three triple interaction terms are positive and statistically significant. Consistent with the hypothesis that social media enables hateful sentiment to spread, a higher reach per AfD user increases the correlation of social media exposure with hate crimes. These interactions work on top of our baseline interaction term, which remains similar in magnitude and highly statistically significant throughout. The smallest coefficient on the triple interaction term of 0.001 in column 3 implies that a one standard deviation increase in likes per user (around 12) increases the baseline coefficient by 25%.[23]

Distracting News Events: As an additional piece of analysis, we investigate the role of news shocks on the transmission of online hate speech to real-world actions. We build on the evidence in Durante and Zhuravskaya (2018), who show that the Israeli army is more likely to strike against Palestinian targets when US media outlets are distracted by other news events. In our case, we hypothesize that other important news events might distract people from the topic of refugees. This is somewhat analogous to Facebook outages in that we exploit additional exogenous weekly variation: if major news events act as a distraction, they should reduce the correlation of exposure to refugee salience with hate crimes.

To measure these news shocks, we obtain Google Trends data on weekly search interest on the terms "Brexit", "Trump", and "UEFA Euro 2016'. Figure A.8 shows that these spike around the respective events. In Table A.15, we show that they are indeed associated with a crowding out of refugee salience: the share of posts about refugees is markedly lower during these key events. As an example, the spike in search interest for Brexit (100 on the Google search index) is associated with an almost 30% drop in the share of refugee posts (relative to the mean).

We next investigate whether, as a result, refugee salience has a weaker link with hate crimes in the weeks these major events attracted particular news attention. If this is the case, we would expect that these events *decrease* the correlation of social media transmission with

[23]To see this, consider that the total implied estimate including interaction is calculated as $0.001 \times 12 \approx 0.012$, which is about 25% than the baseline coefficient of 0.049.

refugee attacks. As before, we implement this by including the Google trends measures as a further interaction in our panel regressions.

Table 5 plots the results. For each of the events in columns 1 to 3, we find a significant negative coefficient on the number of anti-refugee incidents for the triple interaction with distracting news. The negative sign of the coefficient indicates that, during weeks of major news events, changes in anti-refugee incidents correlate less with heightened refugee salience. As the salience of other events crowds that of refugees, there are smaller increases of hate crimes in municipalities with more AfD social media users.

Table 5: News Shock Salience and Hate Crime Propagation

	(1) Brexit	(2) Trump	(3) UEFA EM 2016
AfD users/Pop. × Refugee posts	0.071***	0.096***	0.067***
	(0.018)	(0.022)	(0.017)
AfD users/Pop. × Posts × News shock	-0.019**	-0.009***	-0.002**
	(0.008)	(0.003)	(0.001)
Observations	495,726	495,726	495,726
R-squared	0.078	0.079	0.078
Municipalities	4466	4466	4466
Municipality FE	Yes	Yes	Yes
Week FE	Yes	Yes	Yes

Notes: This table presents the estimated coefficients from a regression of hate crimes against refugees on the interaction of local social media usage and anti-refugee sentiment as in Equation (1). The dependent variable is a dummy for the incidence of a refugee attack. *AfD users/Pop.* is the ratio of people with any activity on the AfD Facebook page to population. *Refugee posts* is the Germany-wide number of posts on the AfD's Facebook wall containing the word refugee ("Flüchtling"). The news shocks refer to the Google searches as indicated in the text. Robust standard errors in all specifications are clustered by municipality. ***, **, and * indicate statistical significance at the 0.01%, 0.05%, and 0.1% levels, respectively.

3.5 Differences Between Social Media And Traditional Media

How does social media differ from traditional media? And could such differences partially explain our results? Existing work has highlighted the ability of users to self-select and interact on social media (e.g. Schmidt et al., 2017). In the following, we highlight three aspects of far-right social media in Germany that may make it a particularly effective transmission mechanism for anti-refugee sentiment compared to mainstream news sources.

First, Figure 7a shows that the share of content about refugees is consistently higher on the AfD's Facebook page compared to traditional news outlets in the Nexis data. The share of refugee mentions on Facebook is also far more volatile and spikes coincide more clearly with salient news events like Merkel's "Wir schaffen das" speech or the Cologne New Year's Eve incidents. In both of these examples, the share of refugee posts on right-wing social media is nearly 100% higher than the share of news stories on refugees, which is consistent with the idea that the topics discussed on Facebook are considerably narrower than in traditional media.

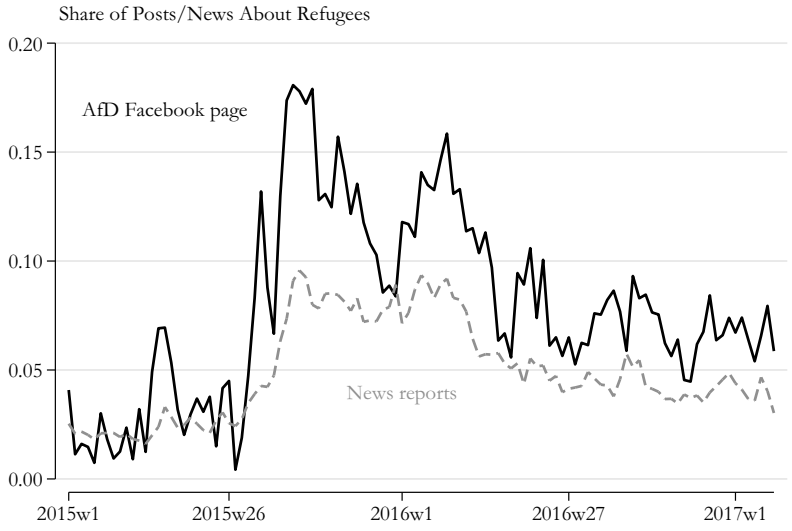
In Figure A.9a in the online appendix, we show that this also holds true in a like-for-like comparison of the share of refugee posts on the AfD's Facebook page relative to the Facebook pages of five major German news outlets. AfD users post twice as much about refugees compared to the next-ranked newspaper. This suggests that the narrowness of content is unlikely to be explained only be the editorial constraints (e.g. space limits in newspapers) of traditional media outlets. Instead, self-selection of like-minded people into the AfD Facebook page likely also play a role. Combined with the interactive nature of social media, this result points towards an anti-refugee group dynamic on the AfD's Facebook page.

Second, as argued by Sunstein (2017), self-selection of like-minded people can lead to the expression of more extreme viewpoints. To shed light on this hypothesis empirically, we compare the full text of news reports about refugees with posts on the AfD Facebook page. Existing reports on far-right hate speech on social media highlight three characteristics as typical (see for example Dinar et al., 2016; Kreiel et al., 2018; Ott and Gür-Seker, 2019): (1) a belief to speak for the "true will" of the people, i.e. the in-group (citizens) compared to the out-group (refugees); (2) an opposition to "elites", in particular politicians and the media, who supposedly mislead or betray the people in an undemocratic way; and (3) a legitimization of discrimination against refugees by highlighting crimes by refugees, an alleged incompatibility of cultural differences, and negative repercussions for vulnerable "locals" (e.g. women, children or pensions).

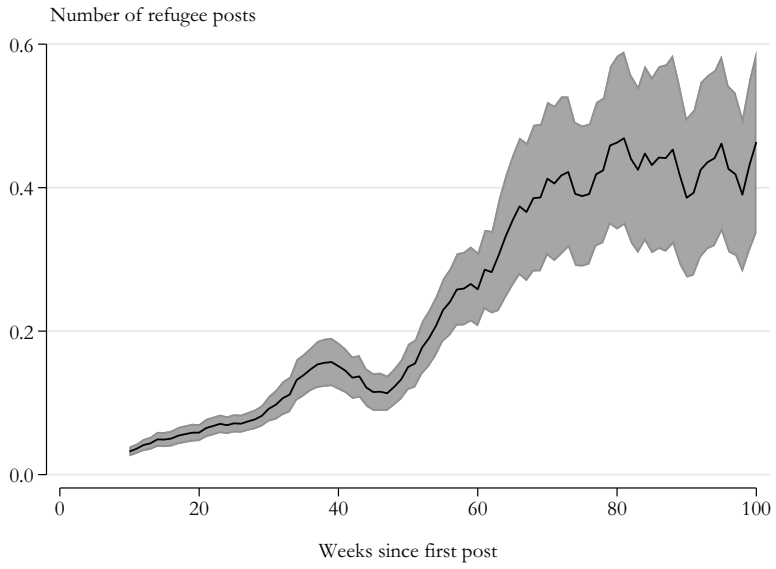
We find evidence for all three of these features of right-wing hate speech on the AfD's

Figure 7: Highlighting Social Media Echo Chambers

(a) Share of Refugee Post over Time



(b) Individual Posting Behavior, by Length of Exposure



Notes: Panel (a) plots the share of posts/reports about refugees on the AfD Facebook page and major German news outlets from Nexis. Panel (b) plots the 10-week moving average of the number of refugee posts per person as a function of a user's time spent on the AfD Facebook page, proxied by the time since the first post. The shaded area indicates 95% confidence intervals.

Facebook page. Our approach is to investigate which words occur with a higher probability in posts on the AfD page relative to news reports in the Lexis corpus.[24] We filter words using the word stems of the German terms for people, elite, democratic, press, crime, foreign, culture, refugee, betrayal, and several vulnerable groups (pensioners, children, women, homeless).

The results of this exercise in Table 6 reveal a clear pattern (see also Table A.16 in the online appendix). As one example, the term "Volksbetrug" (betrayal of the people) is 1715 times more likely to appear on the AfD page than in traditional news outlets. Criticism of "elites" and the media are also far more frequent. Another main difference is how often crimes by refugees are discussed, based on the use of loaded terms like "Flüchtlingskriminalität" (refugee crime). We see expressed fears about "Fremdkulturen" (foreign cultures) and "Burkafrauen" (burka women). This analysis clearly shows that far-right ideas that have widely been interpreted as hate speech are far more pervasive on the AfD page than in traditional media reports.

We find similar results using a text analysis approach using machine learning. In particular, we train a L1 regularized logistic regression model classifier that predicts whether a text comes from the AfD Facebook page or a traditional media outlet. The classifier thereby identifies the words with the highest predictive ability for posts on the AfD Facebook page. Figure A.10 shows a word cloud of the 100 words that best separate social media from traditional media content, based on the model with the highest cross-validated out-of-sample F1 scores.[25] The size of the words represents the magnitude of the coefficients as a measure of variable importance. Consistent with the findings in Table 6, critiques of establishment parties and the economic or social costs of refugees are among the words that most uniquely identify posts on the AfD page.

Third, we investigate how individuals' posting behavior varies with the length of exposure to far-right social media content. We construct a balanced panel of users' activity on the AfD's Facebook page. In Figure 7b, we show users' average number of posts about refugees since their first post on the page. To avoid that a changing sample composition drives our results, we restrict the analysis to the approximately 60% of users who first interacted with the AfD page before June 2015 and thus have been active on it for at least 100 weeks. The results are similar without this restriction.

[24] We calculate word probabilities for each corpus by dividing the number of times a word is mentioned ($Word_i$) by the total number of words in the corpus ($\sum Words_i$), e.g. $P(Word_i^{News}) = Word_i^{News} / \sum Words_i^{News}$. The relative probability is the ratios between the two calculated the two probabilities, i.e. $P(Word_i^{Facebook})/P(Word_i^{News})$.

[25] Note that the model was run in German and the words translated by the authors afterwards. For more details on the machine learning model, see the notes to Figure A.10.

Table 6: Relative Word Frequencies on the AfD Facebook Page

Rank	Word	Translation	Relativ prob.
Panel A: Flücht (refugee)			
1	Flüchtlingsenklaven	refugee enclave	780
2	Flüchtlingslüge	refugee lie	693
3	Flüchtlingsirrsinn	refugee insanity	650
4	Flüchtlingsmafia	refugee mafia	520
5	Flüchtlingsbefürworter	refugee supporter	520
Panel B: Krimi (crime)			
1	Regierungskriminalität	goverment crime	1300
2	Diskriminierungsgesetze	anti-discrimination laws	520
3	Schwerstkriminellen	dangerous criminals	260
4	Fluechtlingskriminalität	refugee crimes	260
5	Kriminalittssteigerung	increase in crime	260
Panel C: Presse (media)			
1	Freie Presse	free press	390
2	Propagandapresse	propaganda press	260
3	Presseempfang	press meeting	260
4	Meinungspresse	opinionated media	260
5	Nazipresse	nazi media	260
Panel D: Volk (people)			
1	Volksbetrug	betrayal of the people	1715
2	volksfeindlich	hostile to the people	780
3	volksverdummenden	brainwashing the people	520
4	Volksverhetzungsparagraphen	law against incitement	520
5	Volksprotesten	protest by the people	260
Panel E: Verrat (betrayal)			
1	Volksverrats	betrayal of the people	130
2	Vaterlandsverrat	betrayal of the fatherland	43
3	Volksverrat	betrayal of the people	43
4	Hochverrat	high treason	36
5	verratenen	betrayed	32

Notes: This table plots the relative probability of words mentioned on the AfD Facebook page compared to reports by major German news outlets on Nexis. We report the results by groups of word stems identified as likely to reflecting right-wing hate speech on social media by previous work in Dinar et al. (2016).

The frequency of refugee posts strongly increases with users' duration on Facebook: within the first year, the average user on the AfD page goes from close to zero to posting at least once about refugees every 2 weeks.[26] This result suggests that the AfD page does not merely attract already active Facebook users with right-wing views, but may increase the willingness of people to express anti-refugee views over time.

This analysis also highlights an important distinction compared to existing research on media and violence. Yanagizawa-Drott (2014) Adena et al. (2015), and DellaVigna et al. (2014) all investigate the effect of nationalistic propaganda in settings of high ethnic tensions. In our setting, there is no nationalistic anti-minority propaganda in traditional media outlets. Rather, we find that social media provides an alternative forum to exchange and spread extreme rhetoric and viewpoints for the fringe elements of society.

3.6 Mechanisms

In theory, multiple mechanisms could be consistent with social media playing a propagating role in real-life hate crimes. We discuss four mechanisms: information exchange, persuasion, collective action, and local spillovers. We provide suggestive evidence that collective action and local spillovers likely play a role in our setting.

First, social media might facilitate the exchange of information. In our setting, relevant information for potential perpetrators could, for example, include the locations of refugee homes and meeting points for demonstrations. We analyze the content of the refugee posts on the AfD Facebook to identify any post that might contain location information. To do so, we tag posts that either contain a zip code, mention the word "straße" (street), "weg" (path), "Flüchtlingsheim", "Asylantenheim", "Flüchtlingsunterkunft" (all three translate to refugee home) or refer to a name of a German town or village.[27] We then manually check the content of tagged posts. This analysis suggests that while some locations like Berlin and Cologne are frequently mentioned in the posts as references to politicians and crimes committed by refugees, we find no mention about specific local information. We found no instance of zip codes or exact addresses. It hence appears unlikely that this channel is the primary driver behind our findings.

A second mechanism could be a persuasion channel, implying that social media persuades potential perpetrators that refugees may be dangerous or undeserving, which may then push

[26] The same holds true for the total number of posts (see Figure A.9b in the Online Appendix).

[27] We base the search on a comprehensive list of 2,061 German towns and 11,000 municipalities from the German statistical office, which covers villages with as little as 20 inhabitants.

some people over the edge. We believe that the timing in our setting makes this channel unlikely. In contrast to other work in Müller and Schwarz (2018) and Bursztyn et al. (2019), we focus entirely on high-frequency variation in social media posts and refugee violence. To the extent that social media changes people's attitudes, this is unlikely to happen in a single week and revert back after anti-refugee salience has subsided. This is particularly true for the results on Facebook and internet outages: it seems unlikely that being cut off from social media during such disruptions reduces hate crimes because potential perpetrators become less xenophobic for a single week.

Third, social media could motivate collective action. Existing evidence in Enikolopov et al. (2016) and Manacorda and Tesei (2016) suggests that social media and mobile internet increase the incidence of protests. In our setting, users could coordinate to carry out hate crimes or learn about others' willingness to carry them out via social media. To investigate this, we rerun the panel regressions in Equation (1) but limit refugee attacks to those undertaken by multiple perpetrators.[28] In line with the collective action hypothesis, Table 7 suggests that our panel regression results are predominantly accounted for by cases with four or more perpetrators. We find no relationship for incidents with fewer than 4 perpetrators. Within the sub-sample where we can identify the number of perpetrators, these attacks account for a similar number of total incidents compared to the cases with more than 4 perpetrators. Hence, this finding is unlikely to be the result of limited statistical power.

Fourth, and somewhat relatedly, it could be that social media enables local spillovers, e.g. through "copy-cat" incidents. This mechanism suggests that potential perpetrators may use social media to learn about other attacks taking place, which could inspire them to carry out additional hate crimes. Because friendship networks on social media are clustered geographically (Bailey et al., 2018), this should be particularly pronounced for attacks happening nearby. We thus again rerun the panel regressions in Equation (1) but now include a dummy variable if neighboring municipalities experience an attack in a given week.[29]

Table A.17 suggests that hate crimes happening in the same week nearby are associated with more anti-refugee incidents. This correlation strongly interacts with the popularity of right-wing social media, particularly when anti-refugee sentiment is elevated. In other words, having an attack in a neighbouring municipality is associated with a stronger correlation of

[28] We were able to hand-code the number of perpetrators for 28% of the hate crimes.

[29] This is akin to the common correlated effects (CCE) estimator proposed by Pesaran (2006) to hold common shocks constant.

Table 7: Mechanism — Anti-Refugee Incidents, by Number of Perpetrators

	(1) Known perp. sample	(2) 1 perp.	(3) <4 perp.	(4) ≥4 perp.
AfD users/Pop. × Refugee posts	0.010**	0.003	0.004	0.007**
	(0.005)	(0.002)	(0.003)	(0.003)
Observations	479,964	479,964	479,964	479,964
R-squared	0.081	0.037	0.046	0.055
Municipalities	4,324	4,324	4,324	4,324
Share of attacks	1	0.245	0.494	0.534
Mean of DV	0.002	0.000	0.001	0.001
Municipality FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes
Baseline controls [8] × Posts	Yes	Yes	Yes	Yes

Notes: This table presents the estimated coefficients from a regression of hate crimes against refugees on the interaction of local social media usage and anti-refugee sentiment as in Equation (1), where we vary the definition of the dependent variable based on the number of perpetrators. All control variables are interacted with the *Refugee posts* measure. Robust standard errors in all specifications are clustered by municipality. ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.1 levels, respectively.

exposure to right-wing social media and the probability of an anti-refugee incident.[30]

Overall, our results appear to be most consistent with the idea that short-run bursts in anti-refugee sentiment on social media can translate into real-life hate crimes by enabling coordination online, both through group actions and local spillovers.

3.7 How Many Refugee Attacks Are Caused By Online Hate Speech?

We conduct a back-of-the-envelope calculation of how many attacks against refugees would have taken place with lower anti-refugee sentiment on right-wing social media. Given that we rely on high-frequency variation, this question is difficult to address. As our estimates are likely to pick up two separate facets of exposure to social media.

On one hand, it could be that exposure to anti-refugee sentiment on social media merely affects the exact timing when refugee attacks occur without changing their total number. On the other hand, the time series of hate crimes and refugee posts on social media in Figure 2 exhibits prolonged overall increases in the number of anti-refugee incidents with the onset of the refugee crisis. These increases are not easy to explain if anti-refugee sentiment exclusively affects the timing of incidents. In our empirical setting, we cannot distinguish between these possibilities.

Despite this important caveat, we still believe it is instructive to assume social media indeed increases the number of hate crimes to illustrate the magnitudes of the results. We calculate the predicted number of attacks, based on the coefficient estimate of 0.024 from a regression with the baseline control variables (see column 1 in Table 2). Multiplying this coefficient with $AfD\ users/Pop.$ and $Refugee\ posts$ gives us the estimated effect on anti-refugee attacks. We sum over all observations to get the total predicted number of anti-refugee attacks as a result of social media. This calculation implies that in absence of social media transmission on social media would result in 289 (10%) fewer anti-refugee incidents.

4 Conclusion

Social media has become a powerful tool for sharing and disseminating information. In this paper, we investigate whether social media can play a role in propagating violent hate crimes. Our findings suggest that social media has not only become a fertile soil for the spread of hateful

[30]Note that, although they are suggestive, we do not interpret these estimates as causal "peer effects", because we cannot distinguish them from common shocks (see Manski, 1993).

ideas but also motivates real-life action. By combining detailed local data on Facebook usage with user-generated content, we can shed light on the link between online posts and anti-refugee incidents in Germany. Plausibly exogenous variation in disruptions to users' Facebook or internet access supports the view that some of the correlations we document reflect a causal effect.

Existing research shows local cultural attitudes towards foreigners are enormously persistent (e.g. Becker and Pascali, 2019; Becker et al., 2016; Voigtlander and Voth, 2012, 2015). We extend this literature by showing that volatile, short-lived bursts in sentiment *within* a given location have substantial effects on people's behavior and that social media may play a role in their propagation. Our findings are particularly timely in light of recent policy debates about whether and how to "regulate" hate speech on social media. Such legislation may come at a high price: since the lines between what constitutes free speech and hate speech can be blurred, regulation can open the door to censorship. Our work does, however, suggest that policymakers ignore online hate speech at their peril. Future research should investigate effective ways to tackle online hate speech. By quantifying the extent of the challenge, our paper takes a first step towards identifying potential harm arising from extended social media usage.

References

Adena, M., Enikolopov, R., Petrova, M., Santarosa, V., and Zhuravskaya, E. (2015). Radio and the Rise of The Nazis in Prewar Germany. *The Quarterly Journal of Economics*, 130(4):1885–1939.

Alesina, A. and La Ferrara, E. (2005). Ethnic Diversity and Economic Performance. *Journal of Economic Literature*, 43:721–761.

Angrist, J. D. and Pischke, J.-S. (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Bailey, M., Cao, R., Kuchler, T., Stroebel, J., and Wong, A. (2018). Social Connectedness: Measurement, Determinants, and Effects. *Journal of Economic Perspectives*, 32(3):259–80.

BAMF (2016). Aktuelle Zahlen zu Asyl. *Bundesamt für Migration und Flüchtlinge*.

Barberá, P. (2014). How Social Media Reduces Mass Political Polarization: Evidence from Germany, Spain, and the US. *Job Market Paper, New York University*, 46.

Bartik, T. J. (1991). *Who Benefits from State and Local Economic Development Policies?* Number wbsle in Books from Upjohn Press. W.E. Upjohn Institute for Employment Research.

BBC (2017). Social Media Warned to Crack Down on Hate Speech.

Becker, S. O. and Pascali, L. (2019). Religion, Division of Labor, and Conflict: Anti-semitism in Germany over 600 Years. *American Economic Review*, 109(5):1764–1804.

Becker, S. O., Pfaff, S., and Rubin, J. (2016). Causes and Consequences of the Protestant Reformation. *Explorations in Economic History*, 62:1 – 25.

Bessi, A., Zollo, F., Del Vicario, M., Scala, A., Caldarelli, G., and Quattrociocchi, W. (2015). Trend of Narratives in the Age of Misinformation. *PLOS ONE*, 10(8):1–16.

Bhuller, M., Havnes, T., Leuven, E., and Mogstad, M. (2013). Broadband Internet: An Information Superhighway to Sex Crime? *Review of Economic Studies*, 80(4):1237–1266.

Boxell, L., Gentzkow, M., and Shapiro, J. M. (2017). Greater Internet Use Is Not Associated With Faster Growth in Political Polarization Among US Demographic Groups. *Proceedings of the National Academy of Sciences*, 114(40):10612–10617.

Bursztyn, L., Cantoni, D., Funk, P., and Yuchtman, N. (2017). Polls, the Press, and Political Participation: The Effects of Anticipated Election Closeness on Voter Turnout. NBER Working Papers 23490, National Bureau of Economic Research, Inc.

Bursztyn, L., Egorov, G., Enikolopov, R., and Petrova, M. (2019). Social Media and Xenophobia: Evidence from Russia. Working Paper 26567, National Bureau of Economic Research.

Card, D. and Dahl, G. B. (2011). Family Violence and Football: The Effect of Unexpected Emotional Cues on Violent Behavior. *The Quarterly Journal of Economics*, 126(1):103–143.

Colussi, T., Isphording, I. E., and Pestel, N. (2016). Minority Salience and Political Extremism.

Dahl, G. and DellaVigna, S. (2009). Does Movie Violence Increase Violent Crime? *The Quarterly Journal of Economics*, 124(2):677–734.

Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., Stanley, H. E., and Quattrociocchi, W. (2016). The Spreading of Misinformation Online. *Proceedings of the National Academy of Sciences*, 113(3):554–559.

DellaVigna, S., Enikolopov, R., Mironova, V., Petrova, M., and Zhuravskaya, E. (2014). Cross-Border Media and Nationalism: Evidence from Serbian Radio in Croatia. *American Economic Journal: Applied Economics*, 6(3):103–32.

DellaVigna, S. and Ferrara, E. L. (2015). Economic and Social Impacts of the Media. NBER Working Papers 21360, National Bureau of Economic Research, Inc.

DellaVigna, S. and Gentzkow, M. (2010). Persuasion: Empirical Evidence. *Annual Review of Economics*, 2(1):643–669.

Dinar, C., Mair, T., Rafael, S., Rathje, J., and Schramm, J. (2016). Hetze gegen Flüchtlinge in Sozialen Medien. *Amadeu Antonio Stiftung*.

Durante, R. and Zhuravskaya, E. (2018). Attack When the World Is Not Watching? US News and the Israeli-Palestinian Conflict. *Journal of Political Economy*, 126(3):1085–1133.

Eisensee, T. and Strömberg, D. (2007). News Droughts, News Floods, and U. S. Disaster Relief. *The Quarterly Journal of Economics*, 122(2):693–728.

Enikolopov, R., Makarin, A., and Petrova, M. (2016). Social Media and Protest Participation: Evidence from Russia. CEPR Discussion Papers 11254, C.E.P.R. Discussion Papers.

Financial Times (2017). Powerhouse Germany Badly Trailing Rivals in Broadband.

Fiorina, M. P. and Abrams, S. J. (2008). Political Polarization in the American Public. *Annual Review of Political Science*, 11:563–588.

Fouka, V. and Voth, H.-J. (2013). Reprisals Remembered: German-Greek Conflict and Car Sales during the Euro Crisis. CEPR Discussion Papers 9704, C.E.P.R. Discussion Papers.

Gabler, N. (2016). The Internet and Social Media Are Increasingly Divisive and Undermining of Democracy. *Alternet*.

Gavazza, A., Nardotto, M., and Valletti, T. (2018). Internet and Politics: Evidence from U.K. Local Elections and Local Government Policies. *The Review of Economic Studies*, 86(5):2092–2135.

Gentzkow, M. (2006). Television and Voter Turnout. *The Quarterly Journal of Economics*, 121(3):931–972.

Hölig, S. and Hasebrink, U. (2016). *Reuters Institute Digital News Survey 2017: Ergebnisse für Deutschland*, volume Nr. 38 of *Arbeitspapiere des Hans-Bredow-Instituts*. Verlag Hans-Bredow-Institut, Hamburg.

Jha, S. (2013). Trade, Institutions, and Ethnic Tolerance: Evidence from South Asia. *American Political Science Review*, 107(4):806–832.

Kreiel, P., Ebner, J., Urban, A., and Guhl, J. (2018). Hass auf Knopfdruck: Rechtsextreme Trollfabriken und das kosystem koordinierter Hasskampagnen im Netz. *Institute for Strategic Dialogue*.

Manacorda, M. and Tesei, A. (2016). Liberation Technology: Mobile Phones and Political Mobilization in Africa. CEPR Discussion Papers 11278, C.E.P.R. Discussion Papers.

Manski, C. F. (1993). Identification of Endogenous Social Effects: The Reflection Problem. *Review of Economic Studies*, 60(3):531–542.

Müller, K. and Schwarz, C. (2018). From Hashtag to Hate Crime: Twitter and Anti-Minority Sentiment. *Available at SSRN: https://ssrn.com/abstract=3149103*.

New York Times (2016). How Facebook Warps Our Worlds, By Frank Bruni.

New York Times (2017a). How Fiction Becomes Fact on Social Media, By Benedict Carey.

New York Times (2017b). Seeking Asylum in Germany, and Finding Hatred, By Ainara Tiefenthler, Shane Oneill and Andrew Michael Ellis .

OECD (2016). Broadband Statistics.

Oksanen, A., Hawdon, J., Holkeri, E., Näsi, M., and Räsänen, P. (2014). Exposure to Online Hate Among Young Social Media Users. *Soul of Society: A Focus on the Lives of Children & Youth*, 18:253–273.

Ott, C. and Gür-Seker, D. (2019). *Rechtspopulismus und Social Media: Wie Wortgebruche in Social Media sprachkritisch betrachtet werden knnen*, pages 279–318. Peter Lang AG.

Pariser, E. (2011). *The Filter Bubble: What the Internet Is Hiding From You*. Penguin UK.

Pesaran, M. H. (2006). Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure. *Econometrica*, 74(4):967–1012.

Pew Research Center (2018). News Use Across Social Media Platforms 2018. Technical report.

Schmidt, A. L., Zollo, F., Del Vicario, M., Bessi, A., Scala, A., Caldarelli, G., Stanley, H. E., and Quattrociocchi, W. (2017). Anatomy of News Consumption on Facebook. *Proceedings of the National Academy of Sciences*, 114(12):3035–3039.

Stephens-Davidowitz, S. (2014). The Cost of Racial Animus on a Black Candidate: Evidence using Google Search Data. *Journal of Public Economics*, 118(C):26–40.

Sunstein, C. R. (2009). *Republic. com 2.0*. Princeton University Press.

Sunstein, C. R. (2017). *# Republic: Divided Democracy in the Age of Social Media*. Princeton University Press.

The Guardian (2017). CPS to Crack Down on Social Media Hate Crime, says Alison Saunders, by Vikram Dodd.

Voigtlander, N. and Voth, H.-J. (2012). Persecution Perpetuated: The Medieval Origins of Anti-Semitic Violence in Nazi Germany. *The Quarterly Journal of Economics*, 127(3):1339–1392.

Voigtlander, N. and Voth, H.-J. (2015). Nazi Indoctrination and Anti-Semitic Beliefs in Germany. *Proceedings of the National Academy of Sciences of the United States of America*, 112(26):7931–7936.

Wooldridge, J. M. (2001). *Econometric Analysis of Cross Section and Panel Data*, volume 1 of *MIT Press Books*. The MIT Press.

Yanagizawa-Drott, D. (2014). Propaganda and Conflict: Evidence from the Rwandan Genocide. *The Quarterly Journal of Economics*, 129(4):1947–1994.

Online Appendix
"Fanning the Flames of Hate: Social Media and Hate Crime"

A A Short History of the AfD

The AfD was founded by Bernd Lucke, a professor of Economics at the University of Hamburg in 2013. Initially, the AfD positioned itself as an opposition party to the common European currency and the bailouts Greece and Spain received as a result of the financial crisis. Right from the start, however, the party also pandered to the right with a conservative social policy. Representatives of the AfD frequently attracted attention for using nationalist terminology and attacking the "Lügenpresse" (Lying Press), a term popularized by the Nazis. With this political program and rhetoric, the AfD attracted 4.7% of the votes in the 2013 German Federal Election, only narrowly missing the 5% electoral threshold.

Nonetheless, the AfD celebrated several victories in state elections and winning seats in the state parliaments of Hesse, Saxony, Thuringia, Brandenburg, Bremen, and Hamburg. Furthermore, the AfD reached 7.1% of the votes in the 2014 European Parliament election. As the Euro Crisis cooled, the party began to shifts its focus further to the right on topics like traditional family values or the role of Islam in Germany. These more nationalist-conservative political positions, championed by Frauke Petry, attracted a significant share of far-right recruits to the party. In 2015, Petry was elected the main speaker of the party, a major defeat for its founder, Bernd Lucke. As a result of this loss, Lucke resigned from his leadership position and left the party completely, followed by several other key party members.

In the run-up to the 2017 federal election, the AfD leadership included Frauke Petry, Alexander Gauland, Björn Höcke, Jörg Meuthen, and Beatrix von Storch, all of whom hold staunch national conservative opinions. With the beginning of the refugee crisis, the aggressively framed mass immigration as dangerous and declared they were unwilling to accept any refugees into Germany. This messaging was accompanied by increased xenophobia and criticism of Islam.

Under the new leadership and impelled by the refugee crisis, the AfD continued to win elections, securing seats in 14 out of the 16 state parliaments in 2016. In the 2017 federal election, the AfD became the third strongest force in the German Parliament with 12.6% of the votes.

B Additional Details on the Data

B.1. A Short Introduction to Facebook Pages and User Data

On Facebook, celebrities, universities, restaurants, and political groups like the AfD have created their own Facebook pages. The AfD page is the starting point for its followers on Facebook. Any Facebook user who is interested in or supports the AfD can "like" its page. The messages posted on the AfD's page then will show up in that user's Facebook feed. The Facebook feed consists of the individualized news and updates every user receives based on his friendship network and interests. In this way, the AfD is able to reach and rally their followers with political messages and party news.

In addition to receiving information from the AfD, Facebook users can become active on the party's page as well. In general, such interactions fall into three categories. First, people can post their own messages, links, or pictures on the fan page. These posts are visible to everybody but will not automatically appear in other users' Facebook feeds. Second, users can comment on posts and comments by other users or the AfD itself. Those comments appear below the original post and are also visible to the public. Third, each post or comment can be "liked" as a sign of support.

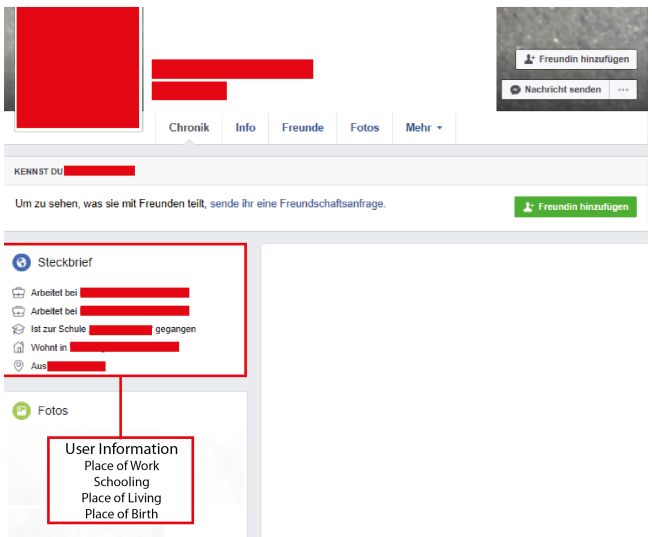
Figure A.1a shows an example of how these three interaction types show up on the AfD page. The Facebook Graph API allowed us to collect all post, comments, and likes from the AfD's fan page, information we highlight ins Figure A.1a. Facebook assigns each user a unique ID that makes it possible to attribute posts and comments to individual profiles.

To hand collect user data, one must visit each individual Facebook user profile, from which, depending on the user's privacy settings, one can determine his/her place of residence and place of birth. Figure A.1b shows an example of a Facebook user profile and where to find the relevant information. If the user decided to hide this information, the box with the user information will be empty.

Figure A.1: Facebook Examples

(a) Alternative for Germany Facebook Page

(b) Facebook User Profile



Notes: Panel (a) shows a screenshot of the Alternative for Germany's Facebook page. The boxes and labels highlight the parts extracted using the Facebook Graph API. Panel (b) shows an example of a Facebook user profile. The box highlights the publicly available user information extracted from Facebook. The authors removed users' personal information for privacy.

Table A.1: Translated Example AFD Posts From Facebook

Date	Post	# Likes	# Comments	# Shares
19/05/2017	Side note in the local newspaper: A Turkish man (23) raped a young woman for more than four hours and was cleared of all charges by the judge. Verdicts that were only known in Arabic cultures are now finding their way into Germany. These pro multi cultural diversity judges are raping the German justice system, cultural sensitivity is apparently more important than the rule of law.	18917	302	1
10/05/2017	+++EUR 204.5 million per month for 500,000 asylum seekers paid in unemployment benefits+++ The top politicians of the old parties sold us the wave of migrants coming to Germany as an enrichment of culture and the economy, but the reality looks very different. The former skilled workers are being financed by social security because they cannot get a job because they are uneducated. Deportations are still not enforced and as a result everybody is fed by the state, even those without asylum.	2418	299	1446
27/12/2016	In Berlin, the police has arrested the wanted teenagers who are suspected to have set a homeless person on fire. Out of the seven suspects between the ages of 15 and 21, six are from Syria and one from Lybia. According to a report of the *Süddeusche Zeitung*, all of them arrived in Germany as "refugees" between 2014 and 2016.	7984	1665	5725
15/11/2016	+ We knew it: More "refugees" are now coming via plane + The government currently has 500 migrants per month flown in from Italy and Greece. The Minister of Interior is further reviewing the admission of an additional 13,500 refugees from Turkey. Only a few European countries are complying with the EU directives, Germany - how could it be any different - is one of the first in line.	2153	1066	2584
21/10/2016	+++ Civil war in Garmisch-Partenkirchen? +++ The *Süddeutsche Zeitung* reports that the situation in Garmisch-Partenkirchen seems to be disastrous: "Blacks have taken over the power in the small skiing village in Germany", the Kremlin-financed Russian station Russia Today reports. The French right-wing news portal Atlantic reports similar things about the alleged regime of dark-skinned refugees and the British Daily Mail speaks of riots in the streets , vandalism, and open sexual assaults.	2084	698	1926

Notes: This table reports five example posts from the AfD Facebook page that were posted by the party itself. The post were translated by the authors.

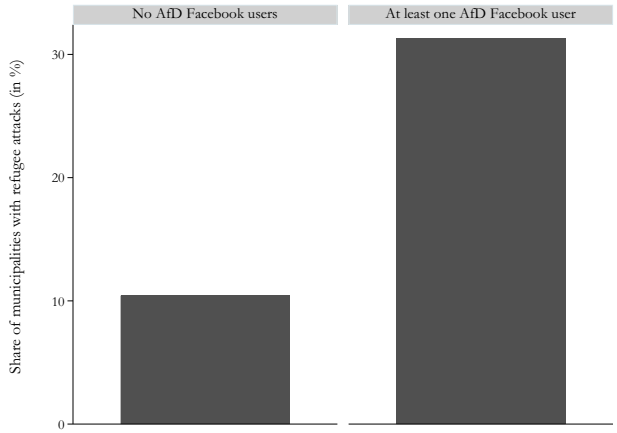
B.2. Examples on Anti-Refugee Incidents

Table A.2: Examples of Anti-Refugee Incidents

Date	Place	Description	Type
03.11.2016	Braunsbedra	A 20-year old Syrian man was riding his bike in the evening and approached a man he assumed needed his help. Suddenly, two additional men appeared and all three started kicking and punching the victim.	Assault
28.12.2016	Langenhagen	An unknown person sprayed graffiti on a refugee home. The graffiti read "Deutsch Nantional (German-National, misspelled in original), "18 (code for Adolf Hitler) and "88 (code for Heil Hitler).	Property Damage
17.11.2016	Oschersleben	A fire occurred in a villa which had until recently accommodated refugees. After a forensic analysis, the police concluded it to be a case of arson, since the fire started in several places at once using fire accelerant. A detonation occurred when the police arrived. Nobody was injured.	Arson
30.01.2016	Schmölln	450 people participated in a demonstration of the "Thügida (Pegida in Thuringia). The police charged 4 people with violating gun control laws and the Public Meetings Act.	Demonstration
30.01.2016	Berlin	The police investigated an insult against inhabitants of a refugee home.	Suspected Cases

Notes: This table reports one example for each class of anti-refugee incidents in the data. The descriptions were translated by the authors.

Figure A.2: Share of Municipalities With Refugee Attacks, by AfD Users



Notes: This figure plots the share of municipalities with at least one refugee attack in our sample by whether we have evidence of at least one AfD Facebook page user in the municipality. We are able to identify one or more AfD users in 3,563 municipalities; for 903 municipalities we find no AfD user.

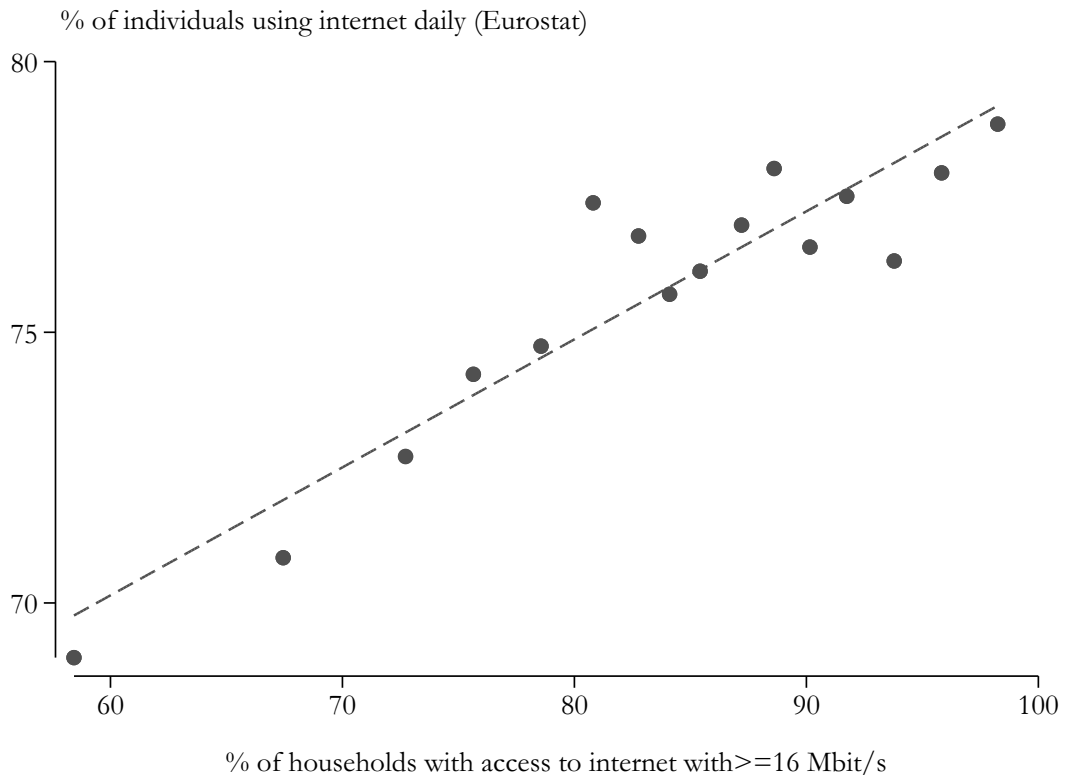
B.3. Additional Variable Overview

Table A.3: Summary Statistics for Additional Controls

	Level	Obs	Mean	SD	Min.	Max.
Additional Media and Internet Controls						
Internet outages/Pop.†	Municipality	495,726	0.007	0.071	0.000	3.614
Registered domains/Pop.	County	495,726	0.141	0.056	0.057	1.390
Mobile broadband speed	County	495,726	12	2	6	24
News paper sales/Pop.	Municipality	491,175	0.092	0.077	0.000	1.644
Additional Right Wing Controls						
Nazi murders/Pop.†	Municipality	495,726	0.002	0.012	0.000	0.282
NPD vote share (2017) (in %)	Election Distr.	495,726	0.491	0.410	0.000	2.006
Ln(1+Deported Jews)	Municipality	495,726	0.606	1.350	0.000	10.930
Ln(1+Stürmer letters)	Municipality	495,726	0.125	0.449	0.000	5.872
Additional Soci-Economic Controls						
Average age	Municipality	479,853	45	2	27	56
Share benefit recipients (in %)	Municipality	495,726	0.424	0.187	0.051	1.207
Share non-Christians	Municipality	479,853	0.325	0.219	0.029	0.941
Manufacturing share (in %)	County	493,062	27	9	2	58
Additional Voting Controls						
CDU vote share (in %)	Municipality	492,618	36	7	20	64
SPD vote share (in %)	Municipality	492,618	19	7	5	47
Left vote share (in %)	Municipality	492,618	8	4	2	26
Green vote share (in %)	Municipality	492,618	7	4	1	25
FDP vote share (in %)	Municipality	492,618	10	3	3	28
Voter turnout (in %)	Election Distr.	495,726	76	3	66	84
Additional Demographic Controls						
Share aged 0-24 (in %)	Municipality	479,853	25	3	14	37
Share aged 25-49 (in %)	Municipality	479,853	33	2	22	45
Share aged 50-74 (in %)	Municipality	479,853	33	3	22	50
Share aged 75 and older (in %)	Municipality	479,853	9	2	4	18

Notes: This table reports summary statistics for the additional control variables in the estimation sample. Variables tagged with a † are scaled by population (in 1,000).

Figure A.3: Daily Internet Users and Share of Households with Broadband Access



Notes: This figure plots the municipal-level share of households with access to broadband internet ($\geq$ 16 Mbit/s) against the state-level percentage of individuals using the internet daily taken from Eurostat survey data, binned into 16 quantiles. The corresponding correlation coefficient is 0.9245.

Table A.4: Overview Variables

(a) Part 1/4

Variable	Level	Description	Source
		Refugee Attacks	
Refugee Attacks/Refugees	Muni.-Week	Constructed by dividing the number of anti-refugee incident in a municipality and week by the number of refugees.	Amadeu Antonio Foundation
Arson Attacks/Refugees	Muni.-Week	Same as Refugee Attacks/Refugees but limited to arson attacks as classified by the Amadeu Antonio Foundation.	Amadeu Antonio Foundation
Other Property Attack/Refugees	Muni.-Week	Same as Refugee Attacks/Refugees but limited to attacks leading to miscellaneous property damages as classified by the Amadeu Antonio Foundation.	Amadeu Antonio Foundation
Assaults/Refugees	Muni.-Week	Same as Refugee Attacks/Refugees but limited to assault as classified by the Amadeu Antonio Foundation.	Amadeu Antonio Foundation
Demonstrations/Refugees	Muni.-Week	Same as Refugee Attacks/Refugees but limited to demonstrations as classified by the Amadeu Antonio Foundation.	Amadeu Antonio Foundation
Suspected Cases/Refugees	Muni.-Week	Same as Refugee Attacks/Refugees but limited to suspected attacks still under investigation, as classified by the Amadeu Antonio Foundation.	Amadeu Antonio Foundation
		Social Media Data	
AfD Users/Pop.	Municipality	The number of AfD Users in each municipality divided by population.	Facebook
Refugee Posts	Week	The number of posts on the AfD Facebook page that contain the word 'Flüchtling' (refugee) in a given week.	Facebook
Posts/AfD Users	Municipality	The total number of posts attributed to AfD users of a given municipality divided by the number of AfD Facebook users.	Facebook
Comments/AfD Users	Municipality	Total number of comments that posts by AfD users of a given municipality received divided by the number of AfD Facebook users.	Facebook
Likes/AfD Users	Municipality	The total number of likes that posts by AfD users in a given municipality received divided by the number of AfD Facebook users.	Facebook

(b) Part 2/4

Variable	Level	Description	Source
		Auxiliary Variables	
$I_{Internet\ Outage}$	Muni.-Week	Dummy variable that is equal to 1 for municipality-week observations that are in the top quartile of the reported internet outages per capita ratio, and 0 otherwise. The number of user-reported outages comes from Heise.de. We exclude outages that are shorter than 24 hours.	Heise.de
$I_{Facebook\ Outage}$	Week	Dummy variable that is equal to 1 for the weeks with major Facebook outages as described in Table A.6, and 0 otherwise.	Various news sources
		Baseline Controls	
Population	Municipality	The population of each municipality in 2015 from the shape file of the BKG. The population numbers in the shape file are equivalent to the 2015 data from the German Statistical Office (Destatis).	BKG/Destatis
GDP/Worker	County	GDP per working population at the county-level.	Destatis
Population density	Municipality	Population density, defined as population over municipality size (in km^2).	Destatis
AfD vote share (2017)	Municipality	The share of votes cast for the AfD in the 2017 German Federal Parliament Election.	Destatis
Share high school	Municipality	The share of people whose highest educational attainment is at least "Abitur", the German high-school certificate.	Destatis
Share Broadband access	County	The share of the population that have access to at least 16 Mbit/s internet connection speed.	BMVI, TÜV Rheinland
Share immigrants	Municipality	The share of the population that are immigrants.	Destatis
Asylum Seekers/Pop.	County	The asylum seekers per capita.	Destatis
		Raw Data	
Refugee attacks	Muni.-Week	The number of anti-refugee incident in a municipality and week.	Amadeu Antonio Foundation
Population (2015)	Municipality	The population for each municipality.	BKG
Refugees (2015)	County	The number of asylum seekers in each county.	Destatis
AfD Users	Municipality	The number of users of the AfD Facebook page we could locate based on their reported place of residence.	Facebook

(c) Part 3/4

Variable	Level	Description	Source
Additional Media and Internet Controls			
Internet outages/Pop.	Municipality	The total number of large internet outages per capita (as defined above) that occurred in a municipality over the sample period	Heise.de
Registered domains/Pop.	County	The number of registered .de domains in a given county, divided by the county population.	Destatis
Mobile Broadband Speed	County	The average mobile download speed in Mbits/s.	BMVI
Newspaper sales/Pop.	Municipality	The number of newspaper copies sold in a given municipality, divided by the population. The data do not contain information for municipalities smaller than 3000 inhabitants, which we impute using the population, population density, AfD vote share, and county fixed effects (results are almost equivalent without imputation).	ZMG
Additional Right-Wing Controls			
Nazi murders (1990-2016)	Municipality	The number of murders classified as having a neo-Nazi motive in a municipality between 1990 and 2016, scaled by population.	Mut gegen rechte Gewalt
NPD vote share (2017)	Election Distr.	The share of votes cast for the extremist right-wing NPD (National Democratic Party of Germany) in the 2017 German Federal Parliament Election.	Bundeswahlleiter
Log(1+Deported Jews)	Municipality	The natural logarithm of the number of Jews who were deported during Nazi times. To analyze cross-sectional correlates, we scale the number of deported Jews by the number of Jews in a municipality in 1933 (see text for details).	Voigtlander and Voth (2012)
Log(1+Stürmer letters)	Municipality	The natural logarithm of the number of letters written to "Der Stürmer", the anti-Semitic newspaper published by Nazi politician Julius Streicher. To analyze cross-sectional correlates, we scale the number of letters by the population in 1933 (see text for details).	Voigtlander and Voth (2012)
Additional Socio-Economic Controls			
Average age	Municipality	The average age in each municipality.	Destatis
Benefit recipients/Pop.	Municipality	The number of social benefit recipients in a given municipality divided by the population.	Destatis
Non-christians/Pop.	Municipality	The number of non-Christians in a given municipality divided by population.	Destatis
Manufacturing share (%)	County	The share of manufacturing employees in a given county.	Destatis

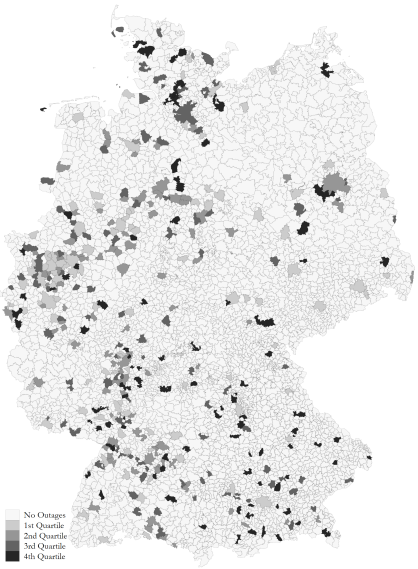
(d) Part 4/4

Variable	Level	Description	Source
Additional Voting Controls (2013 & 2017 Election)			
CDU vote share	Municipality	The share of votes cast for the CDU in the 2017 German Federal Parliament Election.	Destatis
SPD vote share	Municipality	The share of votes cast for the SPD in the 2017 German Federal Parliament Election.	Destatis
Left vote share	Municipality	The share of votes cast for "Die Linke" (The Left) in the 2017 German Federal Parliament Election.	Destatis
Green vote share	Municipality	The share of votes cast for the party "B90/Die Grünen" (Green Party) in the 2017 German Federal Parliament Election.	Destatis
FDP vote share	Municipality	The share of votes cast for the FDP in the 2017 German Federal Parliament Election.	Destatis
Pirate vote share	Municipality	The share of votes cast for the Pirate party in the 2017 German Federal Parliament Election.	Destatis
Voter turnout	Election Distr.	The voter turnout in the 2017 German Federal Parliament Election.	Bundeswahlleiter
Additional Demographic Controls			
Share aged 0-24	Municipality	The number of people aged 0-24, divided by population.	Destatis
Share aged 25-49	Municipality	The number of people aged 25-49, divided by population.	Destatis
Share aged 50-74	Municipality	The number of people aged 50-74, divided by population.	Destatis
Share aged above 75	Municipality	The number of people aged 75 and up, divided by population.	Destatis

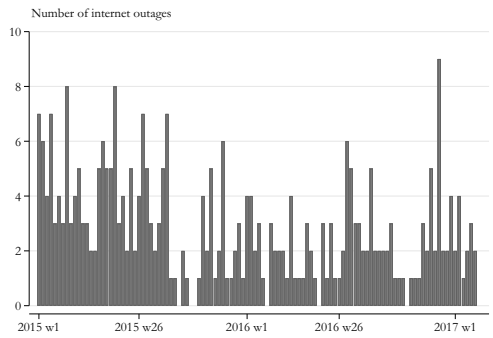
C Additional Details and Results on Internet and Facebook Outages

Figure A.4: Spatial and Temporal Distribution of Internet Outages

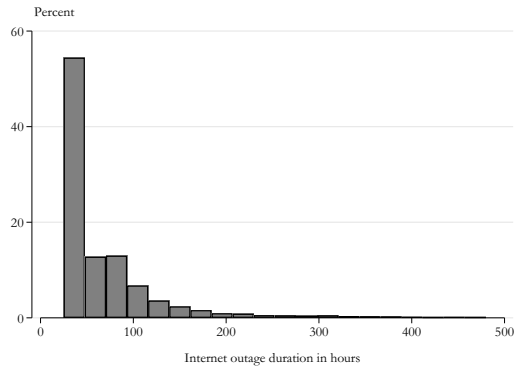
(a) Map of Internet Outages



(b) Internet Outages Over Time



(c) Duration of Internet Outages



Notes: The map in Panel (a) plots the geographic distribution of internet outages per million inhabitants for the German municipalities in the data. Panel (b) plots the distribution of total internet outages per week. Panel (c) plots the distribution of the duration of the individual user reports from *Heise.de* used in the regressions, trimmed at three weeks. See Section 2 for more details.

Table A.5: Validation of Internet Outage Data

(a) Part 1/2

Date	Provider	Region	Description	# Outages	Source
12/06/2015	Kabel D. and Unitym.	Germany	The IT website "Netzwelt" reported a large internet outage on June, 12th 2015. Users of the providers Kabel Deutschland and Unitymedia were especially affected. According to a spokesperson for Kabel Deutschland, the problem was caused by a disruption at the internet hub in Frankfurt.	61	Link
18/06/2015	Unitymedia	Karlsruhe	On July, 18th 2015, the news site "KA News" reported a disruption at internet provider Kabel BW, a subsidiary of Unitymedia. Kabel BW confirmed the problem and explained that their technicians were currently working to fix the problem. The outage affected the area of Karlsruhe.	36	Link
24/06/2015	Unitymedia	Cities in NRW	The "Rheinische Post" reported on June 24th, 2015 that many users of the provider Unitymedia encountered disrupted internet connections beginning on Wednesday, June 23th. Most of the reports came from the cities of Düsseldorf, Mönchengladbach, Neuss, and Münster. Unitymedia did not provide an official statement.	15	Link
05/07/2015	O2 and 1u1	Berlin	The IT website "Golem.de" reported on July, 5th 2015 that users of DSL provider 02 and 1und1 reported disruptions of their internet and phone connections. The problems had started on the June 27th and were largely fixed by the evening of the 5th. Neither provider explained what had caused the problems.	27	Link
08/07/2015	Versatel	Münster	The "Halterner Zeitung" reported on July 8th, 2015 that households in the city of Haltern were cutoff from the internet. The outage was caused by a damaged fiber optic cable. The same cable was also used by internet provider Unitymedia. As a result, Unitymedia users in the Münster area were also affected by the problem.	29	Link
20/08/2015	Unitymedia	NRW and Hessen	The "Gießener Allgemeine" reported on August, 20th 2015 that many users of internet provider Unitymedia encountered disrupted internet connections beginning August 19th. The internet outage affected the entire state of Nordrhein-Westfalen as well as parts of Hessen. At the time of the report, Unitymedia was still investigating the cause of the outage.	81	Link

(b) Part 2/2

Date	Provider	Region	Description	# Outages	Source
04/12/2015	Telekom	Major cities	"ZDnet.de", a website specialized in IT and electronics, reported on December 4th, 2015 that users of the internet provider German Telekom were encountering disrupted internet connections beginning in the early morning of the same day. Most of the reports came from the major cities Berlin, Hamburg, Munich, and Frankfurt. According to the German Telekom, the problem was caused by a breakdown of a RADIUS server that is responsible for authenticating internet access.	19	Link
30/06/2016	Vodafon and Kabel D.	Germany	"Heise.de", the website from which we obtained the outage data, reported on June 30th, 2016 an outage of the internet provider Kabel Deutschland that affected the entirety of Germany. The outage was caused by a problem at a computer cluster. At the time of the report the outage was still ongoing.	122	Link
21/07/2016	T-Online	Germany	The IT website "Golem.de" reported on July ,21th 2016 on internet problems with the provider T-Online. The outage affected not only private households but also business customers of T-Online. A representative of the provider confirmed the problems but did not name any specific cause. At the time of the report technicians were still working to fix the problem.	41	Link
24/11/2016	O2	Germany	On November 24th, 2016 the website "Chip.de" reported a Germany-wide outage of the internet provider O2. The problems were concentrated in metropolitan areas. At the time of report O2 was still investigating the cause of the problem, which was likely an issue with the company's VoIP system.	31	Link
27/11/2016	Telekom	Ruhr area	The "Spiegel" reported on November 27th, 2016 that users of the internet provider German Telekom were cut off from the internet. The outage mainly affected the Ruhr area, but internet problems were also reported in Frankfurt, Hannover, and Braunschweig. At the time of the report Telekom was still working to correct the problem.	19	Link

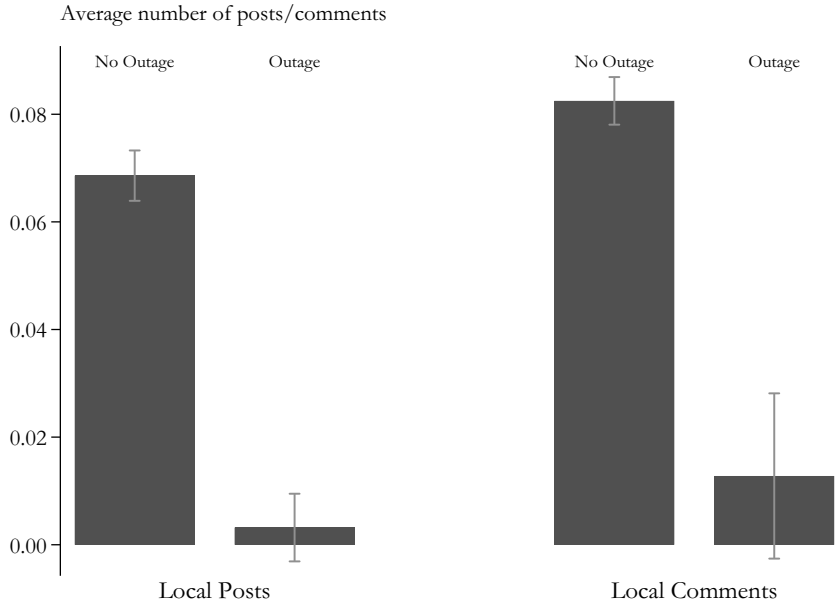
Notes: This table reports several examples of internet outages that were reported in German newspapers or on well-known specialized websites. Each entry lists the date of the outage as well as the affected provider and region. The table further features a short description of the outage and a link to the original news source. The column "# Outage" refers to the number of outages of the affected provider reported by users on Heise.de, which serve as the basis of our internet outage measures; note that this number reflects the number of user reports, *not* the actual number of affected users. The web pages were last accessed in February 2018.

Table A.6: Validation of Facebook Outage Data

Peak Date	Description	Source
26/01/2015	The Facebook page was unavailable globally due to a server error. According to the official statement, the error "occurred after we introduced a change that affected our configuration systems." Initially, the outage had been attributed to an attack by infamous hacker group "Lizard Squad". The outage affected millions of users worldwide, including users of Facebook messenger, Instagram, and the dating app Tinder (which uses Facebook data).	Link
08/04/2015	Facebook users complaint that the site is not loading properly. The outage was particularly concentrated in Western Germany, the Netherlands, and the United Kingdom.	Link
15/07/2015	Facebook suffered a worldwide outage, showing users a simple "Service Unavailable" message. The outage affected all services including the popular Facebook messenger. Although the initial issue was resolved relatively quickly, the problems persisted for many users.	Link
29/09/2015	Users experienced extremely slow or no access to Facebook after a previous disruption on September 24. User reports and news coverage indicate that Germany was particularly badly hit. In a statement to CNBC, Facebook acknowledged the outage and explained that "configuration problems" were at the root of it.	Link, Link 2
14/03/2016	Users in Western Europe - particularly Germany, Austria, Poland, the Netherlands, Belgium, and the United Kingdom - were barred from logging into or commenting on Facebook. The Facebook app was particularly affected.	Link
16/06/2016	Facebook had an outage concentrated in Western Europe. Users were unable to log in, post, use the messenger, or could not access pages (including that of the AfD).	Link
14/09/2016	Worldwide Facebook outage, affecting almost the entire European continent and the eastern United States. Users were unable to log in, post, or read content.	Link
13/01/2017	Users in Western Europe and the eastern United States experienced widespread issues in accessing Facebook, particularly from computer devices.	Link, Link 2

Notes: This table lists the dates of the major Facebook outages that occurred during our sample period. The links lead to the news articles used to identify the disruptions.

Figure A.5: Do Local Internet Outages Reduce Local Facebook Activity?



Notes: This figure plots the arithmetic mean and 95% confidence intervals of local Facebook activity measures based on linking users' locations to their posts and comments. The bars marked "Outage" are municipality-week observations in which a local internet outage occurs. The average values are below one, since we do not observe a post or a comment from each municipality in every week. For example, a mean value of around 0.08 for local comments during weeks without outages implies that on average we observe 1 comment for every 12.5 municipality-weeks pairs (out of 480,963) in our data.

Table A.7: Time Series Evidence — Outages and Aggregate Facebook Activity

	Post outcomes				Outage correlation
	(1) Total posts	(2) Refugee posts	(3) Total posts	(4) Refugee posts	(5) Facebook outage (t)
Facebook outage (t+1),	46.159	-9.413			
	(78.768)	(17.718)			
Facebook outage	-107.372**	-19.880**			
	(50.326)	(7.917)			
Facebook outage (t-1),	-26.517	-15.459*			
	(78.355)	(7.895)			
Internet outage			8.178	0.739	-0.007
			(12.093)	(1.728)	(0.017)
Observations	108	108	109	109	111
R-squared	0.368	0.830	0.344	0.812	0.002
Week-of-year FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes

Notes: This table presents weekly time series regressions of different metrics of Facebook activity on a dummy for Facebook outages in a given week (columns 1-2). Columns 3-5 show that the sum of local internet outages is uncorrelated with Germany-wide Facebook posts and the likelihood of a severe Facebook outage. Newey-West standard errors are reported in parentheses. ** and * indicate statistical significance at the 0.05 and 0.1 level, respectively.

Table A.8: Robustness — Ruling Out Alternative Channels

	(1) Official Reports	(2) Leave One Out Estimator	(3) Lagged Posts	(4) Google Sentiment Measure
AfD users/Pop. × Refugee posts	0.009*	0.057***	0.011	0.103***
	(0.005)	(0.021)	(0.008)	(0.032)
AfD users/Pop. × Posts × Outage	-0.137***	-0.372***	-0.164***	-0.571***
	(0.042)	(0.115)	(0.062)	(0.219)
Observations	474,303	474,303	470,030	474,303
R-squared	0.045	0.084	0.084	0.084
Municipalities	4273	4273	4273	4273
Municipality FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes
All controls [30] × Posts	Yes	Yes	Yes	Yes

Notes: This table presents the estimated coefficients from a regression of hate crimes against refugees on the interaction of local social media usage and anti-refugee sentiment as in Equation (1). The dependent variable is a dummy for the incidence of a refugee attack. *AfD users/Pop.* is the ratio of people with any activity on the AfD Facebook page to population. *Refugee posts* is the Germany-wide number of posts on the AfD's Facebook wall containing the word refugee ("Flüchtling"). Column 1 only uses anti-refugee incidents based on official reports (police or parliament), which are unlikely to be subject to time-varying reporting bias. In column 2 we construct a leave one out measure of *Refugee posts*. Internet outages are defined as municipality-weeks that are in the top quartile of the ratio of reported internet outages to population. Columns 1-3 include all controls as in column 7 of table 2, interacted with *Refugee posts* and all additional interactions of the outage dummy (unreported). Robust standard errors in all specifications are clustered by municipality. ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.1 levels, respectively.

Table A.9: Outage Results with Alternative Standard Errors

Panel A: Internet Outages

	(1) Municipality	(2) Municipality and week	(3) County	(4) County and week	(5) Driscoll-Kraay BW(4)	(6) Driscoll-Kraay BW(10)	(7) Conley 20km	(8) Conley 50km	(9) Conley 100km
AfD users/Pop. × Refugee posts	0.024***	0.024***	0.024***	0.024***	0.024***	0.024***	0.024***	0.024***	0.024***
	(0.009)	(0.009)	(0.009)	(0.008)	(0.009)	(0.009)	(0.005)	(0.005)	(0.006)
AfD users/Pop. × Posts × Outage	-0.181***	-0.181***	-0.181***	-0.181***	-0.181***	-0.181***	-0.181***	-0.181***	-0.181***
	(0.058)	(0.059)	(0.056)	(0.057)	(0.064)	(0.060)	(0.063)	(0.063)	(0.063)

Panel B: Facebook Outages

	(1) Municipality	(2) Municipality and week	(3) County	(4) County and week	(5) Driscoll-Kraay BW(4)	(6) Driscoll-Kraay BW(10)	(7) Conley 20km	(8) Conley 50km	(9) Conley 100km
AfD users/Pop. × Refugee posts	0.027***	0.027***	0.027***	0.027***	0.027***	0.027***	0.027***	0.027***	0.027***
	(0.010)	(0.010)	(0.010)	(0.009)	(0.010)	(0.009)	(0.006)	(0.006)	(0.006)
AfD users/Pop. × Posts × Outage	-0.040*	-0.040*	-0.040*	-0.040*	-0.040*	-0.040*	-0.040**	-0.040**	-0.040**
	(0.021)	(0.021)	(0.023)	(0.023)	(0.024)	(0.023)	(0.020)	(0.020)	(0.017)
Observations	479,964	479,964	479,964	479,964	479,964	479,964	479,964	479,964	479,964
R-squared	0.082	0.082	0.082	0.082	0.082	0.082	0.002	0.002	0.002
Nr. of Clusters	4324	4324/111	394	394/111	4324/111	4324/111			
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline controls [8] × Posts	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table presents the estimated coefficients from a regression of hate crimes against refugees on the interaction of local social media usage and anti-refugee sentiment as in Equation (2). The dependent variable is a dummy for the incidence of a refugee attack. *AfD users/Pop.* is the ratio of people with any activity on the AfD Facebook page to population. *Refugee posts* is the Germany-wide number of posts on the AfD's Facebook wall containing the word refugee("Flüchtling"). Internet outages are defined as municipality-weeks in the top quartile of the internet outage reports to population ratio. Robust standard errors are constructed as defined in the top row. ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.1 levels, respectively.

Table A.10: Outage Results with Alternative Functional Forms

Panel A: Internet Outages

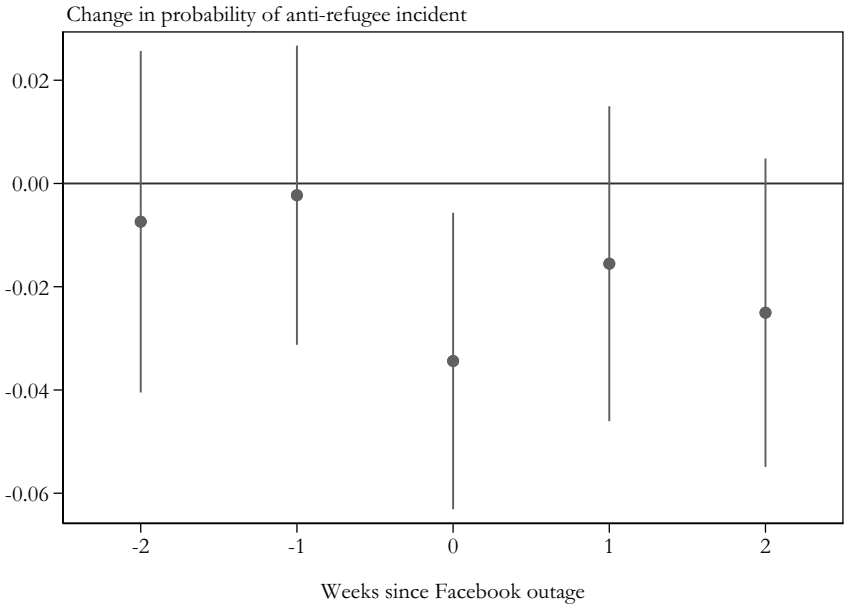
	Refugee Attack Dummy		Refugee Attacks		Log(1 + Refugee Attacks)		Refugee Attacks/Asylum Seeker	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
AfD users/Pop. × Refugee posts		0.016**		0.017*		0.012*		0.018**
		(0.008)		(0.010)		(0.006)		(0.008)
AfD users/Pop. × Posts × Outage		-0.184***		-0.187***		-0.130***		-0.192**
		(0.058)		(0.065)		(0.042)		(0.082)
Outage	-0.004***	-0.002	-0.004***	-0.003	-0.003***	-0.002	-0.002***	-0.002
	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)
Observations	474,303	474,303	474,303	474,303	474,303	474,303	474,303	474,303
R-squared	0.084	0.084	0.157	0.157	0.120	0.120	0.046	0.046
Municipalities	4273	4273	4273	4273	4273	4273	4273	4273
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
All controls [30] × Posts	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Panel B: Facebook Outages

	Refugee Attack Dummy		Refugee Attacks		Log(1 + Refugee Attacks)		Refugee Attacks/Asylum Seeker	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
AfD users/Pop. × Refugee posts		0.021**		0.022**		0.015**		0.022**
		(0.009)		(0.010)		(0.007)		(0.009)
AfD users/Pop. × Posts × Outage		-0.046**		-0.062		-0.036*		-0.039**
		(0.022)		(0.040)		(0.019)		(0.019)
AfD users/Pop. × Outage		1.367		2.700		1.234		0.242
		(1.862)		(3.218)		(1.527)		(1.426)
Outage	-0.001***		-0.002***		-0.001***		-0.002***	
	(0.000)		(0.000)		(0.000)		(0.000)	
Observations	474,303	474,303	474,303	474,303	474,303	474,303	474,303	474,303
R-squared	0.081	0.084	0.155	0.157	0.117	0.120	0.045	0.046
Municipalities	4273	4273	4273	4273	4273	4273	4273	4273
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Week FE		Yes		Yes		Yes		Yes
All controls [30] × Posts	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table presents the estimated coefficients from a regression of hate crimes against refugees on the interaction of local social media usage, anti-refugee sentiment, and outages as in Equation (2). The dependent variable is the measure of refugee attacks listed in the top row. *Refugee Attacks/Asylum Seeker* is the number of anti-refugee incidents per 1000 asylum seekers. *AfD users/Pop.* is the ratio of people with any activity on the AfD Facebook page to population. *Refugee posts* is the Germany-wide number of posts on the AfD's Facebook wall containing the word refugee("Flüchtling"). Note that the interactions of internet outages with refugee posts and AfD Facebook usage are included but unreported to save space. Robust standard errors are constructed as defined in the top row. ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.1 levels, respectively.

Figure A.6: Facebook Outage Event Study



Notes: This figure plots estimates the estimates for λ from an event study regression of Equation (2) which includes 2 leads and lags of the outage interactions. 95% confidence intervals are based on standard errors clustered by municipality.

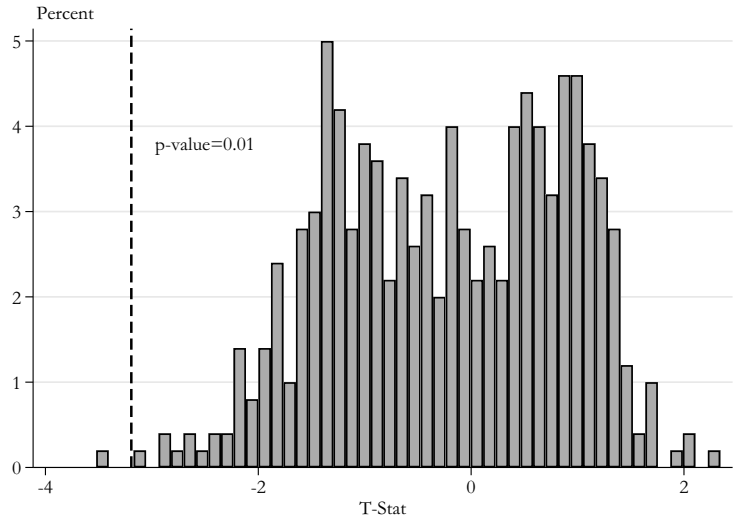
Table A.11: Robustness — Alternative Definitions of Internet Outages

	Change outage definition		**Include shorter outages**		**No extended outages**	
	(1) (Baseline) 75th percentile	(2) 90th percentile	(3) 75th percentile	(4) 90th percentile	(5) 75th percentile	(6) 90th percentile
AfD users/Pop. × Refugee posts	0.024***	0.024***	0.024***	0.024***	0.024***	0.024***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
AfD users/Pop. × Posts × Outage	-0.181***	-0.103**	-0.145***	-0.116***	-0.161***	-0.097*
	(0.058)	(0.052)	(0.043)	(0.040)	(0.058)	(0.055)
Observations	479,964	479,964	479,964	479,964	479,964	479,964
R-squared	0.082	0.082	0.082	0.082	0.082	0.082
Municipalities	4324	4324	4324	4324	4324	4324
Number of outages	308	122	579	231	246	98
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Baseline controls [8] × Posts	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table presents the estimated coefficients from a regression of hate crimes against refugees on the interaction of local social media usage and anti-refugee sentiment as in Equation (2). The dependent variable is a dummy for the incidence of a refugee attack. *AfD users/Pop.* and textitRefugee posts is the Germany-wide number of posts on the AfD's Facebook wall containing the word refugee ("Flüchtling"). Each column includes all additional interactions of the outage dummy (unreported). In columns 1-2, we use the baseline dummy explained above, i.e. outages in the top quartile. In columns 3-4, we include outages shorter than 24 hours (as discussed in Section 2 we exclude this for our baseline measures) and define a new dummy for outages in the top quartile. Columns 5-6 further do not extended outages beyond a single week. Robust standard errors in all specifications are clustered by municipality. ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.1 levels, respectively.

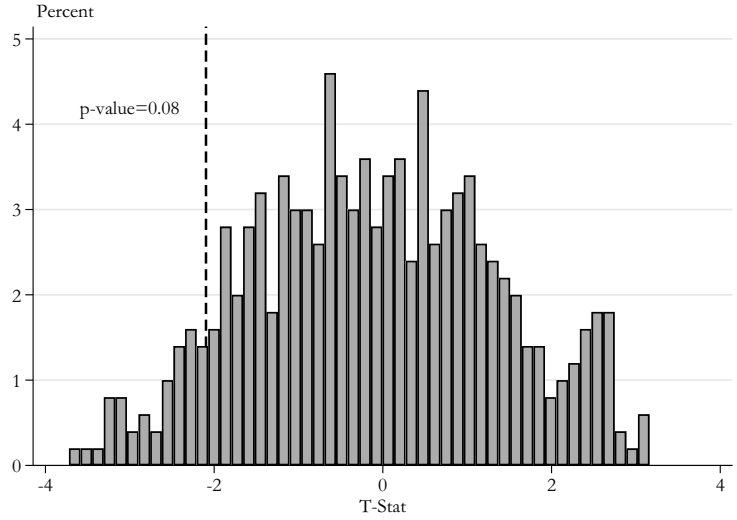
Figure A.7: Randomization Test for Outage Results

(a) Internet Outages



Notes: This figure shows the results of the randomization test, in which we randomly assign placebo internet outages to 313 municipality-week pairs. We repeat this process 500 times and save the t-stat of the triple interaction term of interest. The vertical line marks the t-stat of the actual estimate.

(b) Facebook Outages



Notes: This figure shows the results of the randomization test, in which we randomly assign placebo Facebook outages to eight weeks in our data. We repeat this process 500 times and save the t-stat of the triple interaction term of interest. The vertical line marks the t-stat of the actual estimate.

D Additional Results

Table A.12: Violent vs. Non-Violent Incidents

	(1)	(2)	(3)	(4)
	Arson	Assault	Property Damage	Protest
AfD users/Pop. × Refugee posts	0.002	0.007**	0.012*	0.005*
	(0.002)	(0.004)	(0.006)	(0.003)
Observations	479,964	479,964	479,964	479,964
R-squared	0.016	0.053	0.060	0.060
Municipalities	4324	4324	4324	4324
Municipality FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes
Baseline controls [8] × Posts	Yes	Yes	Yes	Yes

Notes: This table presents the estimated coefficients from a regression of hate crimes against refugees on the interaction of local social media usage and anti-refugee sentiment as in Equation (1). The dependent variable is a dummy for the incidence of a refugee attack; see A.4 for definition of attack types. *AfD users/Pop.* is the ratio of people with any activity on the AfD Facebook page to population. *Refugee posts* is the Germany-wide number of posts on the AfD's Facebook wall containing the word refugee ("Flüchtling"). All control variables are interacted with the $Refugee\ posts$ measure; see text for a description of the controls. Robust standard errors in all specifications are clustered by municipality. ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.1 levels, respectively.

Table A.13: Other Facebook Posts and Anti-Refugee Hate Crimes

	(1) Refugee (Baseline)	(2) Muslim posts	(3) Islam posts	(4) EU posts
AfD users/Pop. × FB posts	4.027***	0.765	-0.043	0.387
	(1.009)	(0.502)	(0.456)	(0.385)
Observations	495,726	495,726	495,726	495,726
R-squared	0.078	0.078	0.078	0.078
Municipalities	4466	4466	4466	4466
Municipality FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes

Notes: This table presents the estimated coefficients from a regression of hate crimes against refugees on the interaction of local social media usage and anti-refugee sentiment as in Equation (1). The dependent variable is a dummy for the incidence of a refugee attack. *AfD users/Pop.* is the ratio of people with any activity on the AfD Facebook page to population. *FB posts* is the Germany-wide number of posts on the AfD's Facebook wall containing the word refugee ("Flüchtling"), with the baseline being *refugee* ("Flüchtling"). Standardized coefficients are reported in square brackets, based on variable transformations with a mean of 0 and a standard deviation of 1. Robust standard errors in all specifications are clustered by municipality. ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.1 levels, respectively.

Table A.14: Social Media Reach and Hate Crime Propagation

	(1) Number of posts	(2) Received comments	(3) Received Likes
AfD users/Pop. × Refugee posts	0.049***	0.048***	0.049***
	(0.017)	(0.017)	(0.017)
AfD users/Pop. × Refugee posts × Reach	0.003***	0.002**	0.001**
	(0.001)	(0.001)	(0.000)
Observations	381,174	381,174	381,174
R-squared	0.086	0.086	0.086
Municipalities	3434	3434	3434
Municipality FE	Yes	Yes	Yes
Week FE	Yes	Yes	Yes
Baseline controls [8] × Posts	Yes	Yes	Yes
Corr(Reach,Population)	0.010	0.011	0.009
Corr(Reach,AfD users/Pop.)	0.017	0.006	0.020

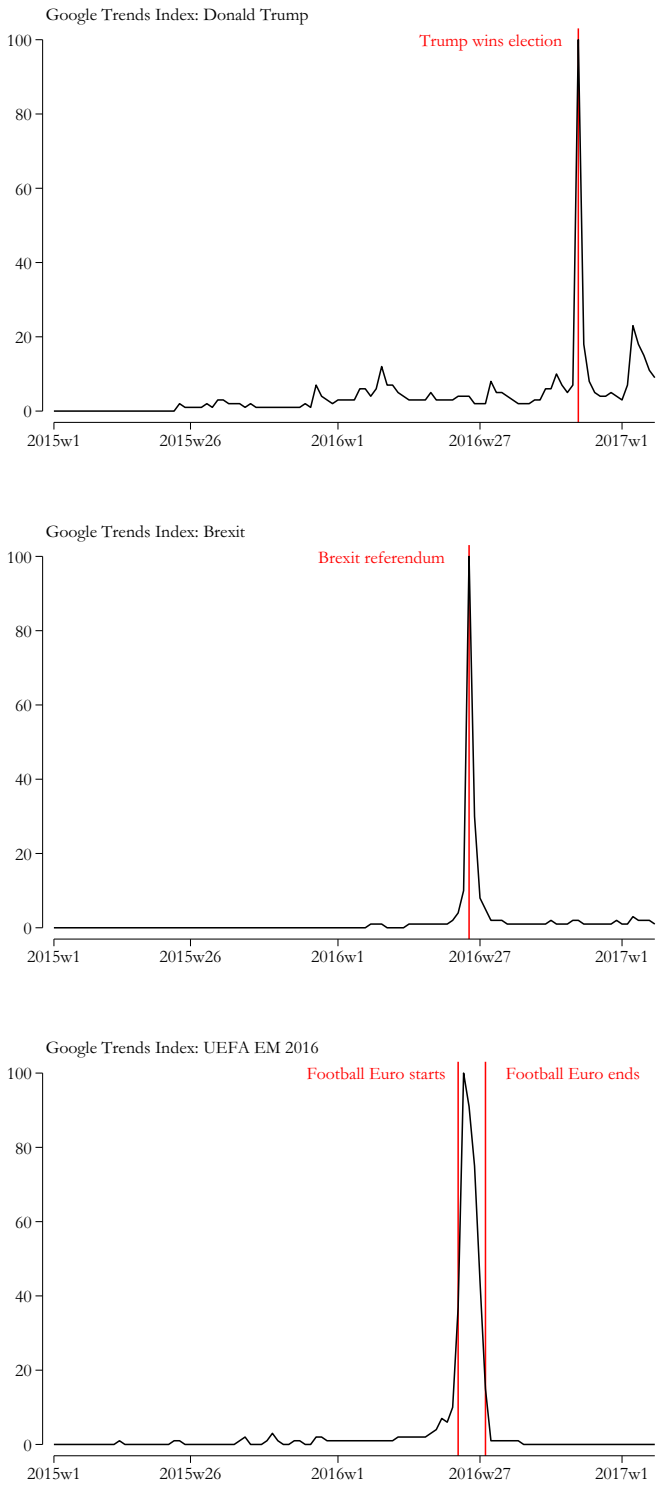
Notes: This table presents the estimated coefficients from a regression of hate crimes against refugees on the interaction of local social media usage and anti-refugee sentiment as in Equation (1). The dependent variable is a dummy for the incidence of a refugee attack. *AfD users/Pop.* is the ratio of people with any activity on the AfD Facebook page to population. *Refugee posts* is the Germany-wide number of posts on the AfD's Facebook wall containing the word refugee ("Flüchtling"). The reach variables in the top row refer to the number of local posts on the AfD wall, as well as comments and likes for AfD posts, all scaled by the number of AfD users (municipalities with zero users are dropped). See text for an explanation of the control variables. Robust standard errors in all specifications are clustered by municipality. ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.1 levels, respectively.

Table A.15: Time Series Evidence – Distractions and Aggregate Facebook Activity

	(1)	(2)	(3)	(4)
	Brexit	Trump	UEFA EM 2016	Horse race
Brexit	-0.023***			-0.041***
	(0.007)			(0.009)
Trump		-0.020***		-0.018**
		(0.007)		(0.008)
UEFA EM 2016			-0.048***	0.058
			(0.013)	(0.043)
Observations	110	110	110	110
R-squared	0.120	0.119	0.098	0.169
Month FE	Yes	Yes	Yes	Yes

Notes: This table presents weekly time series regressions of the share of refugee posts on the AfD page on Google indeces tracking interest in the topics Brexit, Trump, and the European Football Championship. All regressions include week-of-year and month fixed effects. Newey-West standard errors are reported in parentheses. *** and ** indicate statistical significance at the 0.01 and 0.05 level, respectively.

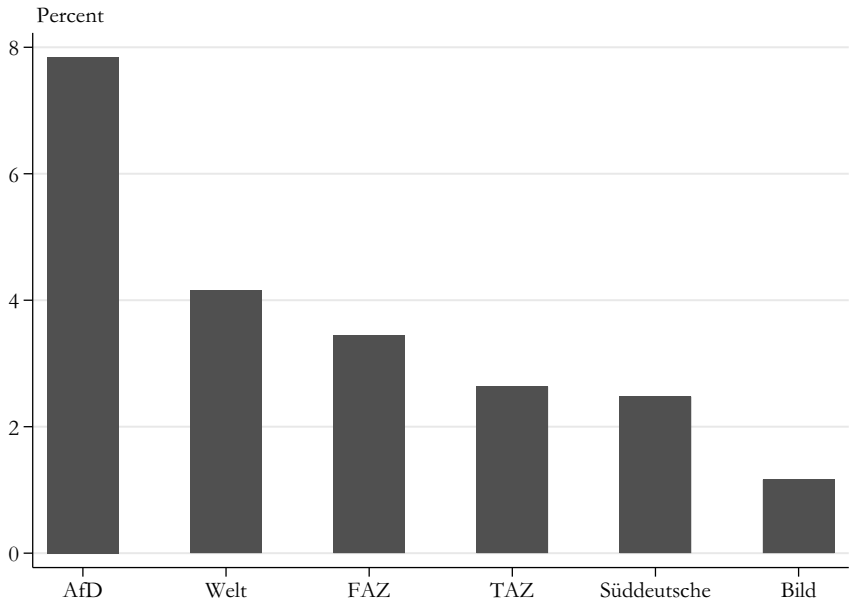
Figure A.8: Google Trends Data — Brexit, Trump, and Football



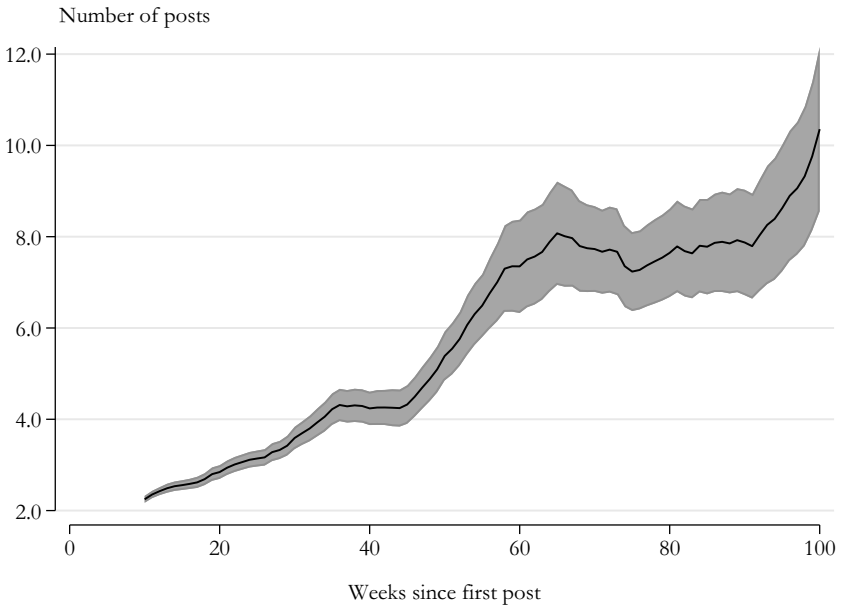
Notes: These figures plot the Google Trends search index (ranging from 0 to 100) for the terms "Trump", "Brexit", and "UEFA EM 2016" for the sample period.

Figure A.9: Differences of Social and Traditional Media — Additional Results

(a) Share of Posts About Refugees on Facebook (in %), by Page



(b) Individual Posting Behavior, by Length of Exposure



Notes: Figure (a) plots the share of posts on the "wall" of the AfD Facebook page and five major German news outlets. Panel (b) plots the 10-week moving average of the number of posts per person as a function of a user's time spent on the AfD Facebook page, proxied by the time since the first post. The shaded area indicates 95% confidence intervals.

Table A.16: Relative Word Frequencies on the AfD Facebook Page

Rank	Word	Translation	Relativ prob.
Panel A: Demokra (democratic)			
1	linksliberaldemokratisch	left-liberal	260
2	Demokratiegedanken	idea of democracy	130
3	postdemokratisches	post-democratic	130
4	scheindemokratische	sham democratic	87
5	pseudodemokratischen	pseudo-democratic	65
Panel B: Renter/Kinder/Frauen/Obdachlose (vulnerable groups)			
1	Kinderbande	gang of children	520
2	Kinderbanden	gangs of children	260
3	Burkafrauen	burka women	260
4	Armutsrentnern	poverty-pensioners	260
5	Kindersex	pedophilia	260
Panel C: Elite (elite)			
1	Elitegruppe	elite group	260
2	Leistungseliten	accomplished elites	130
3	Geldeliten	rich elites	87
4	Staatselite	state elites	87
5	Politelite	political elites	74
Panel D: Fremd (foreign)			
1	Fremdkulturen	foreign cultures	676
2	Fremdvölker	foreign people	260
3	Fremdverwendung	foreign use	260
4	fremdgesteuerten	foreign-controlled	130
5	zweckentfremdeter	misused	130
Panel E: Kultur (culture)			
1	Fremdkulturen	foreign cultures	676
2	Kochkultur	cooking culture	520
3	Kulturgewohnheiten	cultural habits	260
4	Rückkehrkultur	return culture	260
5	Clankulturen	clan culture	260

Notes: This table plots the relative probability of words mentioned on the AfD Facebook page compared to reports by major German news outlets on Nexis. We report the results by groups of word stems identified as likely to reflecting right-wing hate speech on social media by previous work in Dinar et al. (2016).

Figure A.10: Word Cloud — Predictors of AfD Facebook Content

Notes: This figure plots the 100 words with the highest predictive power for AfD Facebook content relative to traditional news media based on an L1 regularized logistic regression model. The size of the words corresponds to the magnitude of the coefficients. The input to the model are the tf-idf weighted words from the AfD's Facebook posts and comments and full text articles from Nexis that contain the word refugee. We exclude words that are shorter than four letters and words that appear in less than five documents. The optimal regularization strength is chosen using 10-fold cross-validation. The selected model achieves an out-of-sample F1 score of above 0.99. The words were translated from German by the authors.

Table A.17: Mechanism — Local Spillovers

	(1)	(2)	(3)
AfD users/Pop. × Refugee posts	0.024***	0.022***	0.016*
	(0.009)	(0.008)	(0.008)
Attack in neighboring municipality	0.004***	-0.000	0.004**
	(0.001)	(0.002)	(0.002)
Attack in neighboring municipality × Posts		0.000	-0.004**
		(0.001)	(0.002)
Attack in neighboring municipality × AfD users/Pop.		13.765***	1.610
		(4.782)	(4.914)
Attack in neighboring municipality × AfD users/Pop. × Posts			0.121**
			(0.052)
Observations	479,964	479,964	479,964
R-squared	0.082	0.082	0.083
Municipalities	4324	4324	4324
Municipality FE	Yes	Yes	Yes
Week FE	Yes	Yes	Yes
Baseline controls [8] × Posts	Yes	Yes	Yes

Notes: This table presents the estimated coefficients from a regression of hate crimes against refugees on the interaction of local social media usage and anti-refugee sentiment as in Equation (1). The dependent variable is a dummy for the incidence of a refugee attack. *AfD users/Pop.* is the ratio of people with any activity on the AfD Facebook page to population. *Refugee posts* is the Germany-wide number of posts on the AfD's Facebook wall containing the word refugee ("Flüchtling"). *Attack in neighboring municipality* is a dummy equal to 1 if a neighboring town experiences a refugee attack in the same week. The coefficient for "Attack in neighboring municipality × Posts" is multiplied by 100 for readability. Robust standard errors in all specifications are clustered by municipality. ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.1 levels, respectively.

E Robustness Checks for Specification

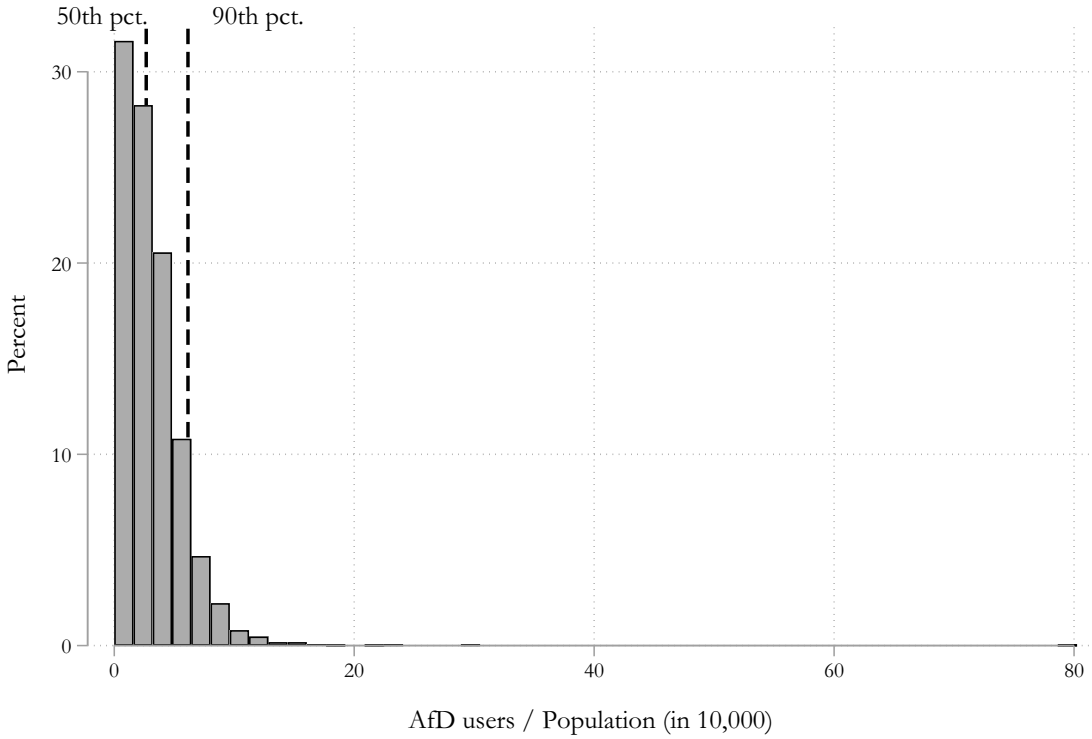
Table A.18: Further Robustness Checks

	(1) Lagged Dependent Variable	(2) Population Weighting	(3) Pre-2015 Users	(4) Refugee Comments	(5) Refugee Post Likes	(6) Refugeee Post Share
AfD users/Pop. × Refugee posts	0.023***	0.029***	0.034**			
	(0.009)	(0.011)	(0.015)			
AfD users/Pop. × Refugee sentiment				0.076***	0.015**	0.437***
				(0.025)	(0.006)	(0.140)
Observations	475,640	479,964	479,964	479,964	479,964	479,964
R-squared	0.085	0.097	0.082	0.082	0.082	0.082
Municipalities	4324	4324	4324	4324	4324	4324
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Baseline controls [8] × Posts	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table presents the estimated coefficients from a regression of hate crimes against refugees on the interaction of local social media usage and anti-refugee sentiment as in Equation (1). *AfD users/Pop.* is the ratio of people with any activity on AfD's Facebook page to population. *Refugee posts* is the Germany-wide number of posts on the AfD's Facebook wall containing the word refugee ("Flüchtling"). Column 1 includes a lagged dependent variable. Column 2 uses weighted least squares (WLS) based on each municipality's population. In column 3, we replace the number of AfD users calculated over the whole sample with the number of users before the sample start (that is, pre-2015). Columns 4 and 5 present results based on the comments and likes (in 100s) of posts on the AfD Facebook page containing the word refugee, rather than the number of posts on the Facebook wall. Column 7 uses the share of posts (in %) containing the word refugee in all posts we observe in a given week. All control variables are interacted with the $Refugee\ posts$ measure; see text for a description of the controls. Robust standard errors in all specifications are clustered by municipality. ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.1 levels, respectively.

E.1. Accounting for the Skewed Distribution of AfD Users

Figure A.11: Distribution of AfD Users / Population



Notes: This figure plots the distribution of the ratio of AfD users in a municipality to population. The vertical lines indicate the 50th and 90th percentile of the distribution, respectively, which we make use of in Table A.19.

Table A.19: Accounting for the Skewed Distribution of AfD Users

	(1) Drop municipalities 0 users	(2) Only above median	(3) Only below median	(4) 10-90th percentile	(5) User quartiles
AfD users/Pop. × Refugee posts	0.053***	0.035*	0.100***	0.076***	
	(0.019)	(0.020)	(0.026)	(0.024)	
AfD users/Pop. (Q2) × Refugee posts					0.012*
					(0.006)
AfD users/Pop. (Q3) × Refugee posts					0.018***
					(0.007)
AfD users/Pop. (Q4) × Refugee posts					0.057***
					(0.011)
Observations	395,493	247,863	247,863	345,876	395,493
R-squared	0.082	0.097	0.026	0.042	0.082
Municipalities	3563	2233	2233	3116	3563
Municipality FE	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes

Notes: This table presents the estimated coefficients from a regression of hate crimes against refugees on the interaction of local social media usage and anti-refugee sentiment as in Equation (1). The dependent variable is a dummy for the incidence of a refugee attack. *AfD users/Pop.* is the ratio of people with any activity on AfD's Facebook page to population. *Refugee posts* is the Germany-wide number of posts on the AfD's Facebook wall containing the word refugee ("Flüchtling"). In column 5, the excluded category is the first quartile of *AfD users/Pop.*; zero-user municipalities are excluded. The coefficients in column 5 are multiplied by 1000 for readability. Robust standard errors in all specifications are clustered by municipality. ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.1 levels, respectively.

E.2. Addressing Many Zeros in the Dependent Variable

A potential concern with the results in the paper is the sparsity of the dependent variable. Even though anti-refugee incidents were relatively frequent in our sample in absolute numbers, they are rare events in the full sample of municipalities and weeks. In this section, we discuss potential issues and provide evidence for the robustness of our estimates.

The first concern is that the standard errors could be biased, which might make it more likely to find statistically significant estimates. We attempt to address this concern by showing that our estimates are robust to a wide array of specification of standard errors (see Table A.9). We also run a randomization check, which yields estimates in line with the p-values from our regressions. Additionally, as shown in Figure 6 and Figure A.6, our regressions also exploit the precise timing of outages. If many zeros would lead us to mechanically find significant effects, this should also be observable in the week before and after the outages, which is not the case in the data. Taken together, the sum of these tests show no evidence that we are overly likely to reject the null hypothesis.

A second concern is that the linearity of the conditional expectation function is violated in our setting and that our coefficients thus misleading. On a basic level, this seems unlikely because the effect of internet outages, for example, is clearly visible in the raw data Figure 5b. To further address these concerns, we re-code the main interaction of interest as the interaction of two indicator variables. In particular, we define periods of "high sentiment" as those in in the top 50% of the weekly number of posts about refugees. We create a dummy for "High Exposure" for towns in the top 50% of the AfD Facebook users to population ratio.

We then estimate our baseline regression models using "dummified" interaction terms.[31] The resulting fully saturated model is completely general and yields unbiased estimates without assumptions about functional form (Wooldridge, 2001, p.456-457): the dummies pick up the mean difference in the number of anti-refugee incidents when refugee salience and exposure to it are high. In fact, these models fit the conditional expectation function perfectly no matter how the dependent variable is distributed (Angrist and Pischke, 2008, p.38). As a result, this approach naturally also accommodates rare events. As we report in Table A.20, our baseline results are essentially unchanged in this specification.

To summarize: we find a correlation between refugee incidents and posts on the AfD page in the time series; a link between these incidents and the usage of the AfD page across

[31] We also create dummies that split each control variable at the median and interact these with the refugee post measure as well as the remaining terms when when including outage interactions.

towns; a panel correlation using the variation in both; an event study effect of Facebook and internet outages; and a breakdown of the baseline panel correlation during outages in the raw data, as visualized in the binned scatter plots above. The panel results also hold in a highly restricted panel where refugee attacks are not rare events and in a fully saturated model that picks up mean differences without assumptions about functional form and is known to fit the CEF perfectly. Despite its caveats, the ZINB results above are also encouraging. Given the range of approaches, we conclude that our findings are highly robust.

Table A.20: Fully Saturated Models

	(1) Panel interaction	(2) Internet outages	(3) Facebook outages
High Exposure × High Sentiment	0.002***	0.002***	0.002***
	(0.000)	(0.000)	(0.000)
Outage		-0.004*	
		(0.002)	
High Exposure × High Sentiment × Outage		-0.009***	-0.003*
		(0.003)	(0.002)
Observations	474,303	474,303	474,303
R-squared	0.082	0.082	0.082
Municipalities	4273	4273	4273
Total attacks	2681	2681	2681
Mean attacks	0.006	0.006	0.006
Municipality FE	Yes	Yes	Yes
Week FE	Yes	Yes	Yes
All controls [8] × Posts	Yes	Yes	Yes

Notes: This table presents the estimated coefficients from a regression of hate crimes against refugees on the interaction of local social media usage and anti-refugee sentiment as in Equation (1). *High Exposure* is a dummy equal to 1 for towns in the top 50% of the *AfD users/Pop.* ratio. *High Sentiment* is a dummy equal to 1 for weeks in the top 50% of *refugee posts*, the Germany-wide number of posts on the AfD's Facebook wall containing the word refugee ("Flüchtling"). The baseline control variables are interacted with the $Refugee\ posts$ measure; see text for a description of the controls. Robust standard errors in all specifications are clustered by municipality. ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.1 levels, respectively.

Table A.21: Addressing Many Zeros in the Dependent Variable

	Panel A: Baseline			
	(1)	(2)	(3)	(4)
	$\geq$ 1 attacks	$\geq$ 2 attacks	$\geq$ 3 attacks	$\geq$ 4 attacks
Refugee posts × AfD users/Pop.	0.194***	0.276***	0.243*	0.305
	(0.051)	(0.092)	(0.138)	(0.203)
	Panel B: Internet outage only			
	(1)	(2)	(3)	(4)
	$\geq$ 1 attacks	$\geq$ 2 attacks	$\geq$ 3 attacks	$\geq$ 4 attacks
Outage	-0.013***	-0.032***	-0.052***	-0.066***
	(0.004)	(0.009)	(0.015)	(0.019)
	Panel C: Internet outage interaction			
	(1)	(2)	(3)	(4)
	$\geq$ 1 attacks	$\geq$ 2 attacks	$\geq$ 3 attacks	$\geq$ 4 attacks
Refugee posts × AfD users/Pop.	0.194***	0.277***	0.243*	0.306
	(0.051)	(0.092)	(0.138)	(0.204)
Outage	-0.011	-0.035*	-0.043	-0.104
	(0.007)	(0.019)	(0.039)	(0.065)
AfD users/Pop. × Posts × Outage	-0.657***	-1.518***	-1.776***	-3.597***
	(0.205)	(0.311)	(0.492)	(1.008)
Observations	136,641	62,715	32,856	20,535
R-squared	0.074	0.090	0.107	0.122
Municipalities	1231	565	296	185
Total attacks	2848	2182	1677	1375
Mean attacks	0.021	0.035	0.051	0.067
Municipality FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes

Notes: This table presents the estimated coefficients from a regression of hate crimes against refugees on the interaction of local social media usage and anti-refugee sentiment as in Equation (1). In each column, we successively restrict the sample to municipalities with a total of at least 1, 2, 3 or 4 attacks on refugees over the sample period. Robust standard errors in all specifications are clustered by municipality. ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.1 levels, respectively.