Public Signal and Private Action: Right-wing Protest and Hate Crimes against Refugees

Sulin Sardoschau* Annalí Casanueva-Artis†

February 2024

Abstract

Citizens can take radical action to advance a political goal and form expectations about the costs and benefits to these actions based on the observed behavior of others. How does a successful protest –a public signal about other people's preferences– influence radical action? We examine this question in the context of the ascent of PEGIDA, Germany's most prominent right-wing movement since World War II. We combine a difference in differences strategy with variation in local weather conditions on scheduled protest days to show that protests on pleasant days attract a higher number of participants, generate more favorable coverage on social media and subsequently cause a surge in hate crimes against refugees. We provide evidence consistent with the notion that local protest and radical action are strategic complements by reducing the perceived social punishment of committing hate crimes (cost channel) and by radicalizing individuals at the margin (preference channel). Protest signals diffuse through right-wing social media networks - but not more general social media networks or geographic proximity. Protest success leads to more overt and brazen forms of violence, kicks off a vicious cycle of violence and generates counter-mobilization in the form of pro-refugee activism online.

Keywords: protest, hate crime, refugees, right-wing
JEL classification: D74, J15, D83, Z10, D72

*Humboldt Universität zu Berlin, School of Business and Economics, and IZA. Email: sulin.sardoschau@hu-berlin.de
†Ifo Institue, Junior Research Group on Social Media and Center for Industrial Organization and New Technologies

1 Introduction

The number of protests worldwide has almost doubled since 1991 (OECD, 2021). This resurgence encompasses a diverse ideological array, spanning from the Arab Spring's liberation movements to civil rights initiatives like Black Lives Matter, and extending to conspiratorial or right-wing factions such as Q-Anon. The efficacy of these movements, particularly in terms of their capacity for mobilization, persuasion, and policy impact, has been extensively scrutinized (Madestam et al., 2013; Sangnier & Zylberberg, 2017; Wasow, 2020). A prominent sub-strand of this literature focuses on the strategic complementarity or substitutability between prospective protest participants (Cantoni et al., 2019; Bursztyn et al., 2021). A critical, yet underexplored, aspect is how the success of a movement - measured by protest attendance and favorable public perception - impacts the likelihood of engaging in radical actions to achieve the same political goal.

We conceptualize radical behavior as a strategic choice between undertaking high-risk, extremist, potentially illegal actions to advance a political agenda or remaining inactive. Individuals form expectations about the costs and benefits of these actions, influenced by the observed behavior of others (Becker, 1968). The cost of radical action consists of a fixed cost (likelihood of detection and punishment) and a social cost (social punishment). Wider public support can serve as a signal about the social cost of radical action and therefore encourage it (Bursztyn et al., 2020). In the extreme, individuals may even derive an added social return, for instance, based on a sense of martyrdom or the desire to spearhead a revolution that is more likely to materialize (Passarelli & Tabellini, 2017). In either case, protest success and radical action act as strategic complements. Conversely, individuals may update their beliefs about the likelihood of reaching the political objective through less costly, established democratic processes, thereby decreasing the relative returns to radical action. In this case, protest success and radical action are strategic substitutes.

We study this question in the context of the largest far-right movement in Germany since World War II: the Patriotic Europeans Against the Islamisation of the Occident (PEGIDA). PEGIDA maintains close ties to neo-Nazis and fascist groups and gained momentum at the peak of the refugee influx to Germany in 2015, mobilizing thousands of protesters across Germany and inspiring offshoots in other countries (Vorländer et al., 2018; Berntzen & Weisskircher, 2016). We investigate whether broader PEGIDA protest participation encouraged hate crimes against refugees and explore the underlying mechanisms, interpreting protest participation as a public signal about support for restrictive refugee policies. Based on this signal, individuals may choose between committing hate crimes to deter immigrants or alternatively pursuing more democratic strategies in the hope of influencing stricter immigration policies through broader public endorsement.

We start by investigating the relationship between right-wing protest participation measured as the log of one plus the number of protest participants and the likelihood of observing a hate crime, using a two-way fixed effects approach that covers the period between 2015 and 2020. Specifically, we estimate a linear probability model and exploit municipality (N=10K) and week (T=251) variation, including a large set of controls that cover time-varying socio-economic, political and technological (i.e. social media) characteristics of these municipalities. In all specifications, we focus on hate-crimes against refugees that were committed in the six days following the protest but not on the protest day itself since we are interested in the public signal of large protests rather than the (potentially violent) dynamics of large protests. We also include lagged hate-crimes and lagged protest participation to account for serial correlation and isolate the effect of large protest on the *onset* of violence against refugees.

This research design has several advantages. First, we are able to control for time-invariant unobserved heterogeneity at the municipality level, capturing - for instance - the root determinants of anti-immigrant sentiment. Similarly, we can account for time varying factors that are common to all municipalities, such as the overall popularity of the right-wing movement, national and European election cycles or the overall salience of the refugee issue.

Our estimates suggest that a 1% increase in the number of participants increases the likelihood of hate crimes by around 0.06 percentage points (pp). This estimate captures the effect of protest participation both at the extensive and intensive margin (i.e. having a protest in the first place and the number of participants). Conditional on observing a PEGIDA protest, we estimate that an increase of 100 participants raises the probability of observing a hate crime by 5pp.

Throughout the paper, we focus on the sub-set of PEGIDA protests that were announced well in advance and always took place on Mondays and exclude protests that erupted spontaneously. Therefore, the probability of observing a protest in the first place is likely orthogonal to time varying and municipality specific factors. We verify this with an event study design, showing that the timing of PEGIDA protests is not correlated with differential trends in the propensity to commit hate crimes.

Nevertheless, the number of participants in a right-wing protest (rather than the protest occurrence itself) could be correlated with unobserved factors that co-determine hate crimes against minorities. For instance, those that intend to commit hate crimes may also be involved in the organisation of the protest and therefore the mobilization of protesters. If such factors vary over time then our estimates would reflect the broader right-wing mobilization potential, rather than a direct causal link between protest attendance and hate crimes.

In order to address this concern, we employ an identification strategy that relies on exogenous variation in local weather conditions at a given protest day. We define a variable that captures pleasant weather, assuming that protests on pleasant days are more likely to attract a higher number of participants (Madestam et al., 2013), draw in a higher number of media outlets or more positive coverage (Zhong & Zhou, 2012) and influence how much participants associate the event with positive sentiments (Goetzmann et al., 2015; Jiang et al., 2022). We verify that moderate temperatures and low precipitation on scheduled protest days are associated with higher protest attendance and a reduction in negative sentiment of PEGIDA content on Twitter.

We estimate the differential effect of a protest during pleasant weather on hate crimes against refugees, focusing on the interaction between a dummy variable for pleasant weather and a scheduled Monday protest. This approach also allows us to improve on previous papers that leverage weather conditions to predict protest participation since we are able to control for protest events and weather conditions separately. Therefore, our specification captures any direct effect of weather on criminal behavior (Blakeslee et al., 2021; Chersich et al., 2019; Michel et al., 2016; Xu et al., 2020) or the direct effect of any protest on hate-crimes. In addition, we condition on the same set of control variables interacted with weather conditions and protest. We include week fixed effects as well as district month of the year fixed effects to account for seasonal weather differences across municipalities, thereby exploiting deviations from average temperatures and precipitation. Consistent with our fixed effects estimation, we find that a protest on a pleasant day increases the likelihood of observing a hate-crime against refugees by around 0.07 pp relative to protests on unpleasant days.

To probe the robustness of our results, we run several empirical checks. First, we assess the problem of forbidden comparisons in two-way fixed effects estimations (De Chaisemartin & d'Haultfoeuille, 2020; Roth et al., 2022). Second, we vary the fixed effect structure (including municipality time trends, state-week fixed effects and more). Third, we account for potential spatial spillovers by using Conley standard errors and using larger geographic units (including county level equivalent to NUTS3 regions and sub-state regions equivalent to NUTS2 regions). Fourth, we run probit and logit estimations. Fifth, we control for weather conditions on the day of the crime. Lastly, we control for the cumulative number of past protest and hate crimes.

Our results show that public support for the PEGIDA movement and radical action act as complements. This can be explained by two forces that are not mutually exclusive and may reinforce one another. On the one hand, public support may be a signal about the social cost of committing a hate crime for an already existing, radicalized fraction of the movement (cost channel). On the other hand, protest success may radicalize new segments of society, increasing the share of people with a preference for committing hate crimes in the first place (preference channel). While we cannot perfectly separate these two channels, we present several pieces of evidence that delineate the scope for both.

We start by investigating the diffusion patterns of the public signal. Individuals will update their beliefs about the preferences of others not only based on local protest but also based on wider public support for the movement. We consider two channels of diffusion: geographic proximity and social media networks. Specifically, we create a measure of right-wing social media connections based on follow relationships between Twitter users and PEGIDA's main account, and tracking tweets and retweets containing the word "PEGIDA". We also account for more general social media connection by creating the equivalent measures based on a random sample of tweets. In a first step, we show that geographic or social media proximity do not predict local protest participation. This may not be surprising since we focus our attention on scheduled protest as opposed to protests that erupt spontaneously (Qin et al., 2021; Potrafke & Roesel, 2022). In a next step, we investigate how the diffusion of this public signal affects radical action. Conditional on local protest size, right-wing social media connections to large protest locations significantly increase local hate crimes, while geographic proximity does not. Overall, our results suggest that large protests as a public signal appear to be more relevant to a subset of individuals that is already sympathetic to the movement, rather than a wider push towards protest participation and hate crimes.

Next, we consider different types of radical action. Assault and rallies, being more visible and direct forms of hate crimes, show a significant increase while the incidence of arson, which can be executed more surreptitiously, does not exhibit a similar spike. This aligns with the hypothesis that individuals feel emboldened to commit acts that are more public and confrontational, suggesting a reduction in the social cost of committing hate crimes.

Conversely, large protests can function as a signal to activists on the opposite political spectrum, raising the relative returns to their radical action. We find that large protest prompt counter-mobilization as evidenced by the increase in the use of the hashtag #refugeeswelcome on Twitter. This suggests that while protests may embolden extremists, they simultaneously inspire actions from those with opposing views, indicating a polarized response to larger protests.

In addition, we analyze the dynamics of belief updating. On the one hand, some theoretical models of radical action conceptualize the primary cause of revolt as an unexpected large shock, typically in the form of a public signal about a surge in grievances within the population (Davies, 1962, 1978; Correa et al., 2021). These "boiling frog models" stand in contrast to the majority of theoretical models, which typically assume that agents engage in Bayesian updating that is independent of the timing and size of the shock (De Mesquita, 2010; Edmond, 2013; Bursztyn et al., 2020). Our analysis refutes the "boiling frog" model of political revolt, showing that past public signals continuously amplify the likelihood of radical actions. Past protest participation and past hate crimes magnify the effect of large protests on hate crimes, suggesting that even small differences in protest size early on can kick-off vicious cycles of violence. In addition, the persistent effect on hate crimes indicates a strategic, rather than merely emotional response by extremists. We show that hate crimes peak a few days after the protest and remain elevated in subsequent two to five weeks.

Lastly, we shed light on local conditions that amplify or mediate hate crimes following large protests. We investigate local social media penetration on Twitter and on Facebook and find that these increase hate crimes, potentially because the public signal is more salient in places that both produce and consume more information about the protest online. We find no evidence that economic grievances, either in the form of natives' or immigrants' economic performance, drives hate crimes.

Our research is closely related with the literature studying the effects of protest on political outcomes. In general, this branch of the literature centers its interest on the impact on electoral outcomes and finds that, in general, social movements tend to increase the vote for parties that defend the same views as the movement (Madestam et al., 2013; Larreboure & González, 2021; Casanueva, 2021; Lagios et al., 2022). Electoral outcomes are crucial to understand changes in political behaviours and preferences of a democratic society. However, electoral outcomes do not capture effects on the very extremes as they are, by definition, bounded within democratic limits. Our paper builds on this literature and adds an analysis of the effects of protest at the extreme right of the political spectrum.

Our paper also adds to the literature on perceived social norms and how public signals, coming from populist entrepreneurs, politicians or election outcomes, may correct these beliefs and influence attitudes and behavior – particularly in populist settings (Bursztyn et al., 2020; Müller & Schwarz, 2021; Bursztyn & Yang, 2022; Müller & Schwarz, 2023; Guriev & Papaioannou, 2022; Ajzenman et al., 2023; Fages & Martínez, 2023). In contrast to these papers, we focus on the role of a single issue, high-frequency, and geographically granular signal about other people's preferences that can complement and substitute radical action in different ways than populist speech or (rare) election outcomes. In addition, the temporal and geographic granularity of the signal allows us to shed light on diffusion mechanisms, which are absent in papers that leverage temporal variation at the national level.

We also contribute to the literature studying the determinants of hate crimes and their consequences. Spikes in hate crimes occur in response to shocks that increase the economic insecurity of individuals (Bray et al., 2022; Jaschke et al., 2022), in response to changes of what a society deems acceptable or desirable (Hanes & Machin, 2014; Romarri, 2020; Carr et al., 2020) or an interaction of both (Entorf & Lange, 2019; Han et al., 2023; Dipoppa et al., 2023). Echoing findings from Müller & Schwarz (2021, 2023); Levy & Mattsson (2023), we demonstrate that online networks play an important role. Social media not only acts as a catalyst for hate crimes but we find that right-wing protests offline serve as a potent precursor, potentially amplifying hate-crime-inciting content on these platforms.

The rest of the paper is organized as follows. Section 2 describes the background and the data. Section 3 explains the different empirical strategies, presents the main results and describes briefly the robustness checks. Section 4 studies the potential underlying mechanisms. Section 5 concludes.

2 Background and Data

2.1 Background

Refugee influx to Germany and hate crimes. Germany has emerged as a primary destination for refugees in Europe, with over 1.6 million asylum applications filed between 2015 and 2018 alone, representing more than 40% of all applications in the European Union during this period (Eurostat, 2019). The surge in asylum applications can be attributed to the eruption of the civil war in Syria and the growing threat of the so-called Islamic State in Iraq, as well as political and social unrest in other parts of the Middle-East and Sub-Saharan Africa leading to a movement of hundreds of thousands of refugees from Syria, Iraq, Afghanistan as well as from Albania, Kosovo and Eritrea.

The peak of asylum applications in Germany occurred in late 2015, following Angela Merkel's controversial decision to admit refugees stranded in Hungary. This decision deviated from the Dublin Regulation, which assigns responsibility for administering an asylum request to the country of first-entry. However, the regulation was effectively abandoned before September 2015, as registration and administrative capacities in Italy and Greece were overwhelmed by immigration pressures, and refugees sought to move to Northern Europe. In March 2016, the European Union established a treaty with Turkey to encourage stricter controls by Turkish authorities at its Western shores, leading to a decline in asylum applications in Germany, which have remained relatively low since then.

In the early stages of the refugee crisis, Germany showed a strong sense of *Willkommenskultur* or "culture of welcome," with many Germans volunteering to help refugees and participating in demonstrations in support of their cause. However, as the number of refugees increased, this sentiment began to shift. Some Germans expressed concerns about the economic and social impact of refugees, with right-wing parties and anti-immigrant groups gaining momentum. The issue became highly politicized, with debates surrounding the government's handling of the crisis and calls for tighter immigration policies. Alongside the success of PEGIDA and the right-wing party AfD, violence against minorities and particularly refugees began to rise and peaked in 2016.

Right-wing protest under PEGIDA. The Patriotic Europeans Against the Islamisation of the Occident, or PEGIDA, movement was founded by Lutz Bachmann in late 2014. It originated in Dresden, the capital of the state of Saxony in Eastern Germany, as a local Facebook initiative with approximately 300 participants in the first demonstration. The movement grew exponentially, following the influx of refugees to Germany in 2015 and reached its peak in mid-January 2015 with approximately 25,000 participants in Dresden. The success was accompanied by offshoots in other cities within and beyond Germany (Berntzen & Weisskircher, 2016).

The movement referenced the renowned Monday demonstrations that took place in the former German Democratic Republic (GDR) in 1989. These demonstrations have since become a symbol of peaceful civic engagement and political change in the minds of many Germans. PEGIDA has appropriated the concept of these demonstrations in an effort to portray itself as a concerned citizens' movement calling for significant reform in immigration policy, as well as the protection of the Christian-Jewish tradition in Europe. Since its inception, PEGIDA has adhered to a consistent three-part structure every Monday, starting with a round of speeches, followed by an evening stroll, and concluding with a closing rally.

Over time, PEGIDA sharpened its profile as a nationalistic, xenophobic and Islamophobic movement with ties to neo-Nazis and other fascist groups (Vorländer et al., 2018). The movement aimed to fuel xenophobic sentiment in the population and gained electoral influence with the rise of a new right-wing populist party - the Alternative for Germany (AfD) in 2016. In 2021, the German domestic intelligence service (*Verfassungsschutz*) has classified the goals of the movement as unconstitutional. Its founder, Lutz Bachmann, was sentenced to two years of probation in 2020 for inciting hate at PEGIDA protests.

The case of Heidenau Heidenau, with a population of 16,000, found itself thrust into the spotlight on August 21, 2015, when local authorities, responding to the escalating refugee crisis in Europe, converted an empty hardware store into a temporary home for asylum seekers. This decision was met with immediate and organized opposition, predominantly fueled by the National Democratic Party of Germany (NPD), known for its right-wing extremist views and for their close ties to PEGIDA. Heidenau is only a 20 minute car ride away from Dresden, PEGIDAs stronghold in Eastern Germany.

At the peak of PEGIDA's success in late summer of 2015, a group of Neo-Nazis rallied around and attacked the refugee camp. The surge in violence led to clashes with the police and sparked a national political debate. The involvement of federal authorities, including the visit by vice chancellor Gabriel and cautionary remarks by chancellor Merkel, meant to emphasize the government's commitment to defending refugee rights and combating extremism. However, these high-profile interventions also sparked controversy, with some critics arguing that the government's response was either too late or insufficiently forceful to deter future xenophobic incidents.

Media outlets rushed into the city to document what was considered "a look into the psyche of the country", quoting – for instance – a witness about seeing "faces that are known here in Heidenau. When the 200 turned up, there were people standing on the railway embankment, who were cheering and clapping. Elderly with bicycles, children were there too. They clapped as if at a summer movie night, as the right-wingers moved towards the Praktiker [hardware store]. And in the sports store, the baseball bats were sold out".[1]

2.2 Main Data Sources

Protest. We take data on PEGIDA protests from the Right-Wing Extremist Mobilization in Germany data set, created by Kanol & Knoesel (2021). The dataset on right-wing extremist demonstrations in Germany between 2005 and June 2020 was created using information from the German federal government's responses to parliamentary questions tabled by the opposition left-wing party *Die Linke*. The dataset includes information on the location, date, number of participants, organizing actors, and mottos of the right-wing extremist demonstrations.

For our purposes, we restrict attention to PEGIDA Monday protests specifically because of their ritualized and pre-planned nature and because these were explicitly anti-refugee. These represent the vast majority of PEGIDA protests (see Figure 1). We focus only on Monday protests between 2015 and the end of 2019, leaving us with a total of 439 protest days with an average of 140 and a maximum of 6,000 participants (see Table 1). Many of the protests were located in the former Eastern part of Germany but we show in Figure 2 that we find offshoots in Western Germany, including the biggest states North-Rhine-Westphalia and Bavaria in the Southern and Western part of the country.

Hate Crimes. We scrape data on hate crimes from the chronicle reported by the Amadeu Anotonio Foundation (AAF) and PRO ASYL Foundation, for the period of 2015-2020. Their data is taken from various sources, including newspaper articles, police press releases, reports from local crime registries and community centers for those affected by right-wing, racist and anti-Semitic violence. Similar to the data on right-wing protests, the most common source are governmental answers to inquiries made by the Left party. Since 2014, every quarter, the Left party in Germany

[1]From one of the main news outlets Die Welt on September 1st 2015.

poses a brief parliamentary question (*Kleine Anfrage*), asking the Federal Government to list all cases of attacks directed at refugees or their accommodation, which are considered by the police as right-wing politically motivated crimes (*PMK*). For each case, the government reports its date, location, and the type of crime committed.

We assign hate crimes to each municipality and week, distinguishing between the weekdays on which these hate crimes were committed. Overall, there are approximately 1,800 week-municipality observations with at least one hate-crime committed against refugees. We show in Figure 2 that the cumulative number of hate-crimes per 100K inhabitants across municipalities between 2015 and 2020 is spread evenly across the country. For the purpose of our analysis, we exclude all hate-crimes that were committed on a Monday to rule out the possibility that the PEGIDA protest itself turns violent or that the crime was committed before the protest. While the former would not invalidate the results, we are more interested in the emboldening effect of large protests rather than the overall probability of an event turning violent when a large number of individuals participates. This restriction reduces the number of municipality-weeks with hate crimes to 448. We show in Figure 1 that the number of hate-crimes per week and the number of protests per week follow similar patterns, with hate-crimes lagging behind by about one to three weeks. This is of course masking a lot of unobserved time-varying heterogeneity, which we will address in the empirical section.

Weather. We take information on weather conditions on protest days from ERA5, which is a global atmospheric reanalysis dataset produced by the European Centre for Medium-Range Weather Forecasts (ECMWF). ERA5 provides a rich historical record of global weather conditions dating back to 1979 from multiple sources, including satellites, radiosondes, and weather stations. ERA5 includes hourly information on a variety of meteorological variables, including temperature, humidity, wind speed, precipitation, and atmospheric pressure, among others. These data are presented at a high spatial and temporal resolution. We extract information on precipitation (rain in mm) and temperature for every hour during protests times, i.e. on Mondays between 12pm and 8pm to create our indicator for pleasant weather. We describe the indicator in more detail in the next section.

Social Media. We use four measures to proxy social media use at the local level: *i*) overall Twitter use, *ii*) PEGIDA tweets, *iii*) followers of the official AfD Facebook account and *iv*) pro-refugee tweets. First, we develop a measure for Twitter usage for each NUTS-3 region in 2013 and 2014 based on a random sample of 600,000 tweets. We geolocate authors using the location indicated in their profile.[2] In addition, we collect all tweets in German and in English containing

[2] We use the Twitter Academic Search API to collect all Twitter data. We pick 6,000 random instants during this period and collect 100 tweets and retweets in German at each instant. Since the Twitter API does not allow to search directly for all tweets in German, we search for tweets containing the 100 most frequent words in German, as listed by Sharoff (2006) on the website `http://corpus.leeds.ac.uk/frqc/`. The Twitter API gives users' location at the time tweets were collected, not posted). We use the Nominatim geocoder from the OpenStreetMaps project to associate the location field to geographical coordinate, and remove locations outisde of Germany, as well as locations that are too general (e.g. "Germany" or "Bavaria").This gives us an estimate of the rate of tweets posted at each instant from each region (expressed as tweets per second), which is then aggregated

the word PEGIDA posted between October 2014 and 2021. This dataset consists of 2,068,258 (and 659,709 geo-localized) tweets and retweets, along with their date of posting, their retweet status, the text of the tweet, and information about the author. We also collect the list of followers of the main PEGIDA Twitter account, and the list of followers of this account. Based on this information we build three indices of social media proximity, which we describe in more detail in section 4.1. Information on the number of followers of the Facebook page of AfD prior to 2015 are taken from Müller & Schwarz (2021). This data is localized at the collective municipality (*Gemeindeverband*) level. These groups may include multiple municipalities. We map the (per capita) number of followers to all municipalities within a group. Lastly, we proxy pro-refugee sentiment, using all tweets and retweets in German mentioning the hashtag #RefugeesWelcome between 2013 and 2018. We are able to geo-localize 150,000 of about 390,000 tweets.

Regional controls Regional controls were retrieved from four administrative sources. Labor market data were taken from the Federal Employment Agency, election outcomes from the Federal Returning Officer (Bundeswahlleiter), and the rest from the Statistical Offices of the Federal States (Statistische Ämter des Bundes und der Länder) and Federal Criminal Police Office (Bundeskriminalamt). Appendix Table C.1 describes all regional controls, their geographic granularity, time coverage and frequency, as well as the data sources. AfD vote share, population density, age structure of population, share of females, share of foreigners unemployed, and share of unemployed are available at the municipality level. Refugee share, share of asylum recipients, share of foreigners with academic qualification, and GDP per population are available at the district-level (*Kreise*). We use NUTS3 boundaries of 2013 to harmonize administrative changes over time, which we describe in more detail in Appendix C.

3 Empirical Strategy and Main Results

3.1 Research Design

Two-way Fixed Effects: Protest Participation and Hate Crimes

We start by investigating the relationship between right-wing protest participation and hate crimes, using a two-way fixed effects approach that covers the period between 2015 and 2020. Specifically, we estimate a linear probability model and exploit municipality (N=10K) and week (T=251) variation to estimate a regression of the following form:

$$\begin{aligned}\text{hate crime}_{it} =& \beta \log(1 + \text{participants})_{it} + \gamma_1 X'_{it} + \zeta_i + \delta_t \\ &+ \gamma_2 \log(1 + \text{participants})_{it-1} + \gamma_3 \text{hate crime}_{it-1} + \epsilon_{it}\end{aligned} \tag{1}$$

Our outcome of interest is a dummy variable for any hate crime against refugees recorded in

at the region-year level.

municipality i in week t. The independent variable is measured as the log number of participants (plus one) at a Monday PEGIDA demonstration in week t . Our coefficient of interest β captures the effect of the number of participants in a right-wing demonstration on hate crimes. X_{it} denotes a large battery of municipality-level time-varying controls, which we detail in Table C.1. We include week and municipality fixed effects, as well as lagged protest participation and hate crime. We cluster the standard errors at the municipality level.

In all specifications, we focus on hate-crimes against refugees that were committed in the six days following the protest but not on the protest day itself since we are interested in the public signal of large protests rather than the (potentially violent) dynamics of large protests. This also alleviates concerns about reverse causality, ruling out the possibility that hate crimes motivate protest participation in the same municipality. Similarly, it is unlikely that hate crimes in one municipality impacts protest participation in all control municipalities. In addition, we restrict our attention to protest participation on scheduled Monday protest and exclude protests that erupted spontaneously to mitigate concerns about unobserved time-varying characteristics at the municipality level that drive both hate crimes and protests. We provide more evidence for the plausibility of this assumption in the next section.

This research design has several advantages. First, we are able to control for time-invariant unobserved heterogeneity at the municipality level, capturing - for instance - the root determinants of anti-immigrant sentiment. Similarly, we can account for time varying factors that are common to all municipalities, such as the overall popularity of the right-wing movement, national and European election cycles or the overall salience of the refugee issue.

In addition, the baseline specification accounts for state-dependence in protest participation and hate crimes. For instance, a protest with a large number of participants may encourage subsequent protest participation. Similarly, hate crimes may normalize future violence against minorities. These dynamic effects may lead to an over-estimation of protest participation on hate crimes. Therefore, we include the lagged number of protest participants as well as the lagged likelihood of observing a hate crime as additional controls.

To account for potential confounding factors and gain precision in our estimates, we include several municipality level time varying controls that proxy the overall economic and political conditions in the municipality, the propensity to commit crimes, as well as the resources or vulnerability of refugees and immigrants more broadly. Hate crimes against minorities require both the presence of minorities and the presence of xenophobic individuals. We account for the presence of minorities by including the share of refugees in a municipality. Similarly, we account for the presence of xenophobic individuals by using the municipality-level vote share for the right-wing party AfD in the latest national or European election. We also control for the overall propensity to commit and record crimes in the municipality with the total number of documented crimes per 100K. In addition, the baseline set of controls includes proxies for time-variant socio-economic conditions in the municipality, i.e. GDP per capita, population density and the share of unemployed.

In a robustness exercise, we complement the baseline set of controls with a large set of additional controls. Specifically, we include the share of unskilled in the population (less than high-school), the share of women, as well as age group dummies (share of population aged 0-25, 25-50, 50-75, and above 75). We account for the vulnerability and resources of the immigrant and refugee community by including the share of tertiary educated foreigners, the share of unemployed foreigners, as well as the share with a granted asylum status. We also include a set of social media controls to account for the possibility that stronger online networks may encourage protest participation and facilitate the coordination of hate-crimes. We include baseline Twitter penetration at the NUTS-3 level in 2014, the number of tweets containing #refugeesWelcome at the NUTS-3 level per capital in 2014 and the number of Facebook users following AfD before 2015 scaled by population at the groups of municipality level.[3]

Recent developments in the literature emphasize caveats in the classical difference in differences setting when it involves both group and time fixed effects, i.e. Two-Way Fixed-Effects or in short TWFE (e.g., De Chaisemartin & d'Haultfoeuille (2020), Wooldridge (2021), Roth et al. (2022), Goodman-Bacon (2021)). We address the issue of forbidden comparisons in the presence of staggered treatment more carefully in Appendix B.1. Let us preview here that the TWFE estimation does not produce any negative weights and is robust to using the difference in differences estimate of De Chaisemartin & D'Haultfoeuille (2022b).

Event Study: First Protest and Hate Crimes

PEGIDA protests were announced and planned well in advance and always took place on Mondays. Therefore, the probability of observing a protest in the first place is likely orthogonal to time varying and municipality specific factors. Nevertheless, it is possible that some municipalities radicalize more quickly than others. These places may also be more likely to mount a PEGIDA protest early on and subsequently attract a higher number of participants. While controlling for lagged protest participation and hate crimes mitigates this problem, it is possible that these dynamics unravelled before the first PEGIDA protest took place and are therefore not captured in the lagged controls.

We leverage an event study design to verify that municipalities with PEGIDA protests are not on different trajectories in terms of their propensity to commit hate crimes. We focus on the very first PEGIDA Monday protest in the municipality and investigate whether time trends in hate crimes differ between treated and non-treated municipalities. We estimate the following specification:

$$\text{hate crime}_{it} = \sum_{k=T_0}^{-1} \beta_k \times \text{protest}_{ik} + \sum_{k=0}^{T_1} \beta_k \times \text{protest}_{ik} + \zeta_i + \delta_t + \gamma X_{it} + \epsilon_{it} \tag{2}$$

We interact time dummies β_k for each week k with a dummy variable $Protest_{ik}$ that switches on after the first PEGIDA Monday protest in the municipality. Our outcome of interest is likelihood of observing a hate crime in municipality i and week t. Equivalent to our baseline specification, we

[3]These controls are measured at baseline and therefore captured by the fixed effects. However, in the next section we will leverage the interaction between these controls and time-varying factors and we are currently expanding the set of social media controls to account for changes in social media use over time.

include municipality ζ_i and week fixed effects δ_t, as well as the same set of baseline controls X_{it}. We cluster standard errors at the municipal level.

To assess the parallel trends assumption, we consider three different control groups in this event study (Miller, 2023). First, we run a naive event study including never-treated and later treated municipalities in the control group. However, municipalities that never experience a PEGIDA protest may be selected on unobservables and not provide a good counterfactual, potentially overestimating the effect on hate crimes. Therefore, in a second step, we only include municipalities that have ever been treated in the observation period. This relaxes the identifying assumption in the sense that control municipalities with no PEGIDA protest are not the counter-factual for over time change, but that among municipalities that ever experience a PEGIDA protest, the timing of this protest is as good as random.[4]

Lastly, we only include municipalities with at least one recorded hate crime against refugees. This changes the counterfactual to municipalities with a similar propensity to commit hate crimes. If there were unobserved time-varying factors that drive hate crimes across all "radical" (i.e. prone to commit hate crimes) municipalities but only ignite protests in some, we would observe a correlation between protest timing and hate crimes for the full sample. However, restricting attention to radical municipalities reveals whether these react to the same unobserved shock. For instance, a political scandal about refugees may incite violence in all radical municipalities but only ignite protest in the most populous ones. If all radical municipalities react to the same event, then the coefficients β_k for the post protest period should not be significantly different across these municipalities.

Figure 3 plots the β_k coefficients for the ten weeks leading up to and after the first PEGIDA protest. Sub-figure a shows the naive event study design, sub-figure b shows the results for the ever-treated sample, and sub-figure c shows the results for the sample of municipalities with hate crimes or "radical" municipalities. Across all three samples, we find no evidence of different pre-trends. In addition, we find that PEGIDA protest is associated with a rise in hate-crimes in the following two to five weeks. Reassuringly, there is no evidence that radical municipalities react to the same unobserved shock.

Quasi-Experimental Approach: Weather Conditions at Monday Protest

In the previous section, we have provided some evidence that the occurrence of PEGIDA protest is plausibly orthogonal to time-varying unobserved factors at the municipality level. However, we are also interested in the extensive margin of the effect of PEGIDA protest in the sense that attendance and public perception are a signal about the social cost of committing hate crimes. While scheduled protests themselves may not be endogenous, the number of participants in a right-wing protest could be correlated with unobserved factors that co-determine hate crimes against minorities. For instance, those that intend to commit hate crimes may also be involved in the organisation of the protest and therefore the mobilization of protesters. In the case that these variables change in time, β would not capture the causal effect of protest participation on hate crimes but the far right

[4] In Appendix Figure A.2, we show that there is heterogeneity in treatment timing.

mobilization potential both in the form of encouraging protest participation and hate crimes.

In order to address this concern, we employ an identification strategy that relies on exogenous variation in local weather conditions at a given protest day. We define a variable that captures pleasant weather, assuming that individuals at the margin are more likely to join a Monday protest if weather conditions are good. We show in Appendix Figure A.3 that protest participation follows an inverse U-shape, indicating that more protesters take to the street in moderate temperatures and that protest participation decreases with higher levels of precipitation. We identify the appropriate weather conditions for protest participation in Appendix Figure A.4. We report the coefficients for separate regressions that estimate the effect of different rain and temperature cut-offs interacted with the protest dummy on the log number of participants to show that rain above 10 mm and temperatures above 21 degrees Celsius are associated with significant drops in the number of protest participants.[5] We combine these two components into a dummy variable for pleasant weather during protest times (between 12 pm and 8 pm) that switches on if temperatures are between 0 and 21 degrees Celsius and there is no heavy shower (average precipitation of less than 10 mm per square metre).

We estimate the differential effect of a protest during pleasant weather on hate crimes against refugees. This approach also allows us to control for protest events and weather conditions separately. In addition, we condition on the same set of control variables interacted with weather conditions and protest. We include week fixed effects as well as district month of the year fixed effects to account for seasonal weather differences across municipalities, thereby exploiting deviations from average temperatures and precipitation. We estimate the following reduced form regression:

$$
\begin{aligned}
\text{hate crime}_{it} = \; & \beta \text{ protest}_{it} \times \text{weather}_{it} \\
& + \eta_1 \text{ protest}_{it} \times X'_{it} \\
& + \eta_2 \text{ weather}_{it} \times X'_{it} \\
& + \gamma_1 X'_{it} + \gamma_2 \text{ protest}_{it} + \gamma_3 \text{ weather}_{it} \\
& + \mu_{im} + \delta_t + \epsilon_{it}
\end{aligned}
\tag{3}
$$

A causal interpretation of β requires that the interaction between weather and protest only impacts hate crimes through its effect on protest participation and other factors related to increased public attention for right-wing protests. For instance, protests on pleasant days may attract relatively more journalists or generate more positive imagery of the protest. This does not invalidate our empirical strategy since we are interested in the emboldening effect of seemingly successful right-wing protest. This can be in the form of protest participation, in related positive protest coverage (Zhong & Zhou, 2012) or in a increase in the positive mood of people when participating in or learning about the protest (Goetzmann et al., 2015; Jiang et al., 2022). We test this hypothesis in

[5] In these regressions we include again municipality-month of the year fixed effects as well as week fixed effects, control for protest and weather cut-off separately and include the interaction between the full set of controls and protest and weather cut offs.

Appendix Figure A.5, where we estimate specification 3 but use the share of PEGIDA tweets with negative sentiment as our outcome of interest.[6] We find that following protest on pleasant days the negative sentiment significantly decreases in the same week (and in the following three weeks).

Exploiting exogenous variation in the interaction between protest events and pleasant weather allows us to address several endogeneity concerns. First, determinants of hate-crimes that are related to any right-wing protest itself will be captured in η_1. For instance, the most extreme fraction of the right-wing movement that participates in every scheduled protest (irrespective of weather conditions) could use the occasion to conspire and coordinate hate crimes against minorities. Second, any weather conditions that are conducive to crimes more generally and hate-crimes more specifically, will be captured in η_2 (Heilmann et al., 2021; Field, 1992). In addition, controlling for weather conditions (and their interaction with municipality characteristics) separately alleviates many of the concerns associated with a violation of the exclusion restriction in the context of weather data (Sarsons, 2015; Mellon, 2021).

Crucially, we control for district month of the year fixed effects, which allow us to interpret our weather variable as weather shock and not average weather conditions. This rules out the concern that the scheduling of protests in a certain district is based on expected weather conditions, which could also be related to district-specific seasonal differences in the propensity to commit hate crimes. One could imagine that - for instance - some districts that are more religious than others would not schedule a PEGIDA protest during Christmas season. These may also be less likely to commit hate crimes when the salience of the church and its welcoming approach to refugees is high. Lastly, it is unlikely that very local deviations from average weather conditions in a given week, impact protest participation in all other municipalities.

In Figure 4, we present a plausibility check for the assumption that protest during pleasant weather is exogenous to unobserved municipality characteristics. Specifically, we predict the probability of treatment with an array of municipality-level characteristics. These characteristics are standardized and categorized into demographics (age and gender distributions), socioeconomic status (employment, education, and income metrics), political preferences (party vote shares and voter turnout), cultural/religious composition (shares of religious affiliations), migration-related factors (asylum, refugee, and migration shares), and crime statistics.[7] The coefficients demonstrate no systematic correlation with the treatment, reinforcing the plausibility of our identifying assumption. Importantly, the municipality fixed effects and controls capture all characteristics presented in Figure 4 and others that could be related to selection into treatment.

[6] Tweets are geo-located and assigned to the NUTS3 region. Sentiment analysis is conducted using the machine learning tool BERTA based on a German language dictionary.

[7] The model, described by $\sum Pleasant \times Protest_{muni.year} = \beta X_{muni.year} + \sum Pleasant_{muni.year} + \sum Protest_{muni.year} + \gamma_{state} + \theta_{year} + \epsilon_{muni.year}$, regresses a comprehensive set of municipal characteristics against the count of protests on pleasant days per municipality and year. State and year fixed effects are incorporated, along with controls for the number of pleasant days and Monday protests. Standard errors are clustered at the state level.

3.2 Main Results

In this section, we present evidence on the relationship between the prominence of PEGIDA's Monday protests and the incidence of hate crimes. In Panel A of Table 2, we examine the impact of the number of protest participants, logged to address non-linearity, on the probability of observing a hate crime in the same municipality during the week of the protest. We exclude hate crimes on the protest day to focus on the signaling effect to extremist fractions rather than potentially violent dynamics of large events. Panel B presents the alternative reduced form approach, assessing the influence of favorable weather conditions during a protest on the likelihood of subsequent hate crimes. This serves as a proxy for protest salience, reflecting participation and engagement levels.

In Panel A, our findings consistently indicate a positive and statistically significant correlation between the PEGIDA protest participation and the occurrence of hate crimes. Importantly, the introduction of increasingly comprehensive control variables (described in the previous section) across columns 1 to 4 in both panels does not significantly change the effect size or alters the explanatory power of the model but it substantially reduces the number of observations.[8] If anything, the precision of our estimates increases across columns. To maximize the number of observations and circumvent the problem of bad controls, we consider the restricted set of controls in column 1 as our preferred specification but note that the results hold throughout the paper when including the more extensive set of controls.

The effects of protest attendance on hate crimes is large. In column 1 of Panel A, we find that a 1% increase in protest attendance is associated with a 0.06 percentage point rise in the likelihood of hate crimes. This effect captures protest at the intensive and extensive margin. When restricting attention to municipalities that have hosted a protest, a one standard deviation increase in the number of participants (approximately 330 participants) is linked to a 16.5 percentage point increase in hate crime probability in the following days.

In Panel B, we focus on the reduced form results. As mentioned above, we control for weather and protest occurrence separately and successively interact both with the set of baseline, demographic, refugee and social media controls in columns 1 to 4. The presence of pleasant weather during a protest—indicative of higher engagement—increases the probability of a hate crime by approximately 0.07 percentage points compared to protests occurring in less favorable weather. This specification uniquely captures the intensive margin of the effect, as it conditions on the occurrence of a protest. Similar to Panel A, the inclusion of detailed controls, including interactions with weather conditions, does not meaningfully alter the magnitude or significance of our findings. Overall, our analysis underscores the pivotal role of protest visibility and engagement in influencing social outcomes, particularly hate crimes, within affected municipalities.

[8]Note that the social media controls in Panel A do not vary over time and are thus captured by the fixed effects. However, in Panel B, social media controls are interacted with time-varying protest dummies and weather conditions to account for the differential effect of both depending on baseline social media use.

3.3 Robustness Checks

We conduct several robustness checks, detailed in Appendix B. As mentioned above, the address recent developments in the literature on TWFE (De Chaisemartin & d'Haultfoeuille, 2020; Wooldridge, 2021; Roth et al., 2022; Goodman-Bacon, 2021). Specifically, we address the concern of "forbidden comparisons" in the presence of staggered treatment for our baseline estimation and our quasi experimental approach. We present these results in more detail in Appendix B.1, but let us preview here that the share of negative weights is almost zero for our baseline specification but increases for the quasi-experimental approach as we add a more extensive set of controls (those beyond baseline controls). In order to avoid these forbidden comparisons and considering the robustness of our coefficients to adding more controls, we focus on the baseline specification presented in Panel A column 1 of Table 2.

Additionally, we run several empirical checks presented in Table B.2 to B.4. We first show that the results stay similar when varying the fixed effect structure (including municipality time trends, state-week fixed effects and more). Then, we show that the results are robust to changing the geographical aggregation to NUTS-3 and NUTS-2; to restricting the sample to municipalities that had at least one protest; controlling weather condition on the day where hate crimes where committed and controlling for the cumulative number of past protest and hate crimes. We change the estimation method and show that the results are robust to Probit and Logit estimations instead of linear probability models. Finally, we estimate our standard errors allowing for spatial correlation at different distance windows and present the results in Table B.4.

4 Mechanisms

Our results show that public support for the PEGIDA movement and radical action act as complements. This can be explained by two forces that are not mutually exclusive and may reinforce one another. On the one hand, public support may be a signal about the social cost of committing a hate crime for an already existing, radicalized fraction of the movement (cost channel). On the other hand, protest success may radicalize new segments of society, increasing the share of people with a preference for committing hate crimes in the first place (preference channel). Overall, it is very difficult to disentangle the two mechanisms and they may also not be separable conceptually.[9] Nevertheless, we present several pieces of evidence that delineate the scope for both mechanisms, starting with diffusion patterns of protest signals before moving on to the dynamics of belief updating. We also provide additional evidence on substitution towards more overt hate crimes, counter-mobilization of pro-refugee activists and discuss mitigating forces, including economic grievances and social media use.

[9] See Battaglini et al. (2021) for a theoretical discussion of how information aggregation and the coordination motives in protests are conceptually intertwined. These insights extend to a framework where radical action is another form of political participation with a reduction in cost and a change in the information set of individuals.

4.1 Diffusion of the public signal

Right-wing networks on social media and geographic proximity. Our results show that protest success in a municipality increases hate crimes in the same municipality. However, if protest success is a public signal it may diffuse across space and embolden extremists elsewhere. One way in which information about protests spreads is through social media networks (Qin et al., 2021). Conditional on local protest size, social media connections to other municipalities with large protests could therefore amplify the effect on hate crimes.

To test this hypothesis, we consider PEGIDA networks on Twitter. We develop a measure of social media proximity N between municipality i and municipality j based on retweets of tweets containing the word PEGIDA. For each retweet, we locate the original user and the retweeting user. To compute the time-varying influence from NUTS-3 region j to i, we count the number of retweets in the last 6 months and divide by the population of the retweeting NUTS-3 region.In addition, we leverage the follower relationship between NUTS-3 regions for a constant measure for PEGIDA networks. More specifically, we scrape the followers of the main PEGIDA Twitter account and their respective followers. The influence of region j on region i is defined by the number of geo-localized Twitter users (specifically PEGIDA followers) in NUTS3 region i following users in region j.

Moreover, we create the equivalent measure for Twitter networks more broadly, drawing from the random sample of 600,000 tweets described in section 2. This allows us to distinguish right-wing PEGIDA networks from generic digital proximity. Since we do not have the follower relationships at baseline for this sample, we use the retweets before 2015 to create the time-constant measure of social media proximity. We define S^{Γ}_{it} as the social media proximity-weighted sum of protest participants in all other municipalities in week t, where Γ can either be the PEGIDA network P or the wider Twitter network W. It is worth noting that social media proximity can be asymmetric. In other words, the influence of municipality j on i can be larger than the influence of municipality i on j, if Twitter users in i are more likely to follow and retweet Twitter users from j.

$$\mathrm{S}^{\Gamma}_{it} = \sum_{j \neq i} \mathrm{N}^{\Gamma}_{j \rightarrow i(t)} \times \log(1 + \text{participants})_{jt}$$

We are interested in the estimating the effect of S^{Γ}_{it} - the Twitter connections to municipalities with large protests - on hate crimes in municipality i and week t. Our most stringent specification writes as follows:

$$\begin{aligned} \text{hate crime}_{it} = \quad & \beta_1 \mathrm{S}^{P}_{it} + \ \beta_2 \mathrm{S}^{W}_{it} + \beta_3 \mathrm{G}_{it} \\ & + \beta_4 \log(1 + \text{participants})_{it} \\ & + \gamma_1 \log(1 + \text{participants})_{it-1} + \gamma_2 \text{hate crime}_{it-1} \\ & + \gamma_3 X'_{it} + \zeta_i + \delta_t + \epsilon_{it} \end{aligned} \tag{4}$$

We face three challenges in isolating the effect of β_1 on hate crimes. First, social media networks could just be a proxy for geographic proximity. If there are spatial spillovers of protests and hate crimes, we may falsely attribute this to social media networks. This is why we include a geographic proximity measure G_{it}, which is the the distance weighted sum of protest participants in other municipalities. Specifically, we compute the influence of municipality i to municipality j as the distance between the centers of the two municipalities, applying a linearly decreasing window function until 100 km distance. Including both geographic proximity, wider social media networks and right-wing networks also allows us to distinguish between large protest as a signal to the broader public versus large protest as a signal to a sympathetic audience, i.e. PEGIDA tweeters and followers.

Second, we face a simultaneity problem. Digitally connected municipalities could experience a surge in protest attendance that is inspired by protest attendance elsewhere. In this case, we would not strictly identify the effect of social media networks but that of local protest attendance driven by social media networks. This is unlikely because we measure protest on the same day and protest mobilization is unlikely to happen in real time. Nevertheless, we address this concern in two ways: for one, we can directly measure the effect of social media networks on protest attendance on the same day and on the following Monday. In addition, when looking at the effect on hate crimes, we always condition on local protest attendance and therefore measure the additional effect of being connected to places with a large number of participants.

Lastly, digitally connected municipalities may react to the same shock that drives hate crimes and protest participation. We address this concern by investigating the persistence of the effect, i.e. wee look at the effect of S_{it} on hate crimes in t as well as $t+1$. If extremists make an update their beliefs about the preferences of others, this effect should persist over time. If digitally municipalities react to the same time-varying factor, we may not observe the same level of persistence. Importantly, we simultaneously control for local protest attendance in t, as well as past protest attendance and hate crimes in $t-1$. This means that we capture deviations from local trends in protest attendance and hate crimes, isolating the marginal effect of protest attendance in digitally connected municipalities.

We present the results in Table 3. Panel A uses our time-varying measure of social media proximity based on retweets in the six months preceding the protest. Panel B uses the constant social media network based on the location of PEGIDA followers and retweet relationship of the wider social media network at baseline. In all specifications, we include both the standardized social media proximity S_{it}^{Γ} and the geographic proximity G_{it} weighted distance to protest locations. Therefore, our coefficients can be interpreted as the effect of a one standard deviation increase in (geographic or social media) connectedness on the likelihood of observing a protest (column 1) or hate crime (columns 4 and 5) and on the percentage change in protest participation (columns 2 and 3).

Spillovers to protest. In column right-wing social media proximity S_{it}^{P}, we find positive but small and noisily estimated effects on protest occurrence and size , both for time-varying (Panel A) or constant (Panel B) measures of Twitter networks. We see a positive and small effect on the likelihood of observing a protest and the size of the protest suggesting that protest may diffuse through right-wing social media networks. This is either because individuals sympathetic to the movement are more likely to mount and participate in protests themselves, or because they are able to mobilize others if they see more people taking to the streets elsewhere. However, the effect becomes even smaller and vanishes when we use baseline social media proximity ion Panel B. Turning to wider social media proximity to large protest locations S_{it}^{W}, we also do not find any effect on the likelihood of PEGIDA protests or the number of participants.

In addition, we find no evidence that geographic proximity to other protest locations G_{it} increases protest probability and size. If anything, protests in geographic proximity decrease protest mobilization, potentially through substitution away from one location to another. This is also the case when we apply a more stringent distance cut-off at 50km or 25km (not reported).[10] Throughout, geographic and social media spill-overs to other municipalities do not play a significant role in our context. This may not be surprising since we focus our attention on scheduled protest on the same day as opposed to protests that erupt spontaneously which have shown to be more responsive to nearby protests in other settings (Qin et al., 2021; Potrafke & Roesel, 2022).

Spillovers to hate crime. We investigate the effect of social media and geographic proximity to municipalities with large protest on the likelihood of observing a hate crime locally in columns 4 and 5. We condition on local protest participation to isolate the additional effect of connectedness to other protest locations on hate crimes. Throughout, local protest participation continues to play the most important role in determining the onset of hate-crimes in a municipality. However, right-wing social media connections strongly predict hate crimes above and beyond local protest participation, both for our time-varying measure of social media proximity in Panel A and the time constant measure at baseline in Panel B. Again, geographic proximity to large protest locations does not predict hate crimes, which also alleviates concerns about spatial spillovers in the treatment. In addition, wider social media proximity does not matter in the short run. In contrast, wider social media proximity at baseline predict hate crimes later on.

Overall, these findings are consistent with several explanations. For one, the public signal of large protests may be more relevant to some groups than to others. Geographic proximity and wider social media proximity are more crude measures of signal strength, reaching a broader audience. If this signal moved the entire distribution of preferences towards hate crimes, we may still expect to see an increase in hate crimes on average. Instead, the public signal appears to be more relevant to a subset of individuals that is already sympathetic to the movement, i.e. those that actively engage with PEGIDA on social media. In addition, the long-run effects of wider social media connectedness suggests that users connected to large protest locations may radicalize and

[10]We verify in a robustness check that our results hold when we expand the unit of analysis from municipalities, to counties to sub-sate regions.

join the PEGIDA social media network over time. Moreover, our findings are also consistent with the literature on overconfidence, extremism and correlation neglect (Ortoleva & Snowberg, 2015), where radicalization can stem from the inability to recognize that much of the information on social media comes from people similar to themselves. Individuals in digitally connected municipalities may over-interpret the signal of large protests elsewhere and engage in extremist behavior as a consequence.

4.2 Types of radical action

Overt versus covert hate crimes. Protest success can serve as a public signal about social cost associated with hate crimes. Some hate crimes can be committed without revealing ones identity while others are more overt, objecting the perpetrator to more social punishment. If extremists consider protest participation as a public signal that normalizes hate crimes, we expect that they are more ready and willing to commit overt crimes rather than covert ones. Our data set allows us to distinguish between the type of hate crimes, including assault, rallies, arson and other crimes. In the left panel of Figure in the left panel of Figure 5, we illustrate the coefficients from our TWFE regression at the top (participants) and the reduced form regression at the bottom (protest on pleasant days). We separately estimate the likelihood of observing a hate crime in each of these categories, including all other crimes and no hate crimes in the control group. This allows us to detect potential substitution effects from one crime category to the other.

Our analysis reveals a discernible pattern in the types of hate crimes committed following protests. Assault and rallies, being more visible and direct forms of hate crimes, show a significant increase, aligning with the hypothesis that individuals feel emboldened to commit acts that are more public and confrontational. In contrast, the incidence of arson, which can be executed more surreptitiously, does not exhibit a similar spike and even decreases in the reduced form specification. This distinction suggests that the public nature of the protests serves as a catalyst for more brazen forms of hate crimes, where perpetrators feel a reduced social barrier or heightened justification for their actions.

Counter-mobilization and pro-refugee activism. Protest success is not only a signal to individuals sympathetic to the movement but also to those opposed to it. Pro-refugee activists may equally update their beliefs about the relative returns to radical action and adjust their behavior accordingly. We have shown in the previous section that large protests do not have geographic spill-over effects but diffuse through right-wing networks, suggesting that the signal may not generate complementarities in the wider population but for a specific subset of individuals for which refugee issues are particularly salient.

We test this hypothesis by looking at pro-refugee activism online. Specifically, we scrape the universe of tweets using the hashtag #refugeesWelcome between 2015 and 2018.[11] This hashtag was a rallying cry for activists that supported a humane policy towards refugees. We verify that

[11]Due to Twitter API access limitations we cannot expand this data set to 2020.

these tweets capture pro-refugee attitudes by employing a machine learning sentiment classification neural network from Guhr et al. (2020) and by training an algorithm with 1,000 manually coded tweets. In Table A.1, we replicate our main table but use the per capita number of pro-refugee tweets at the NUTS3 level during the week of the protest as our outcome of interest. We show TWFE and reduced form estimates in Panel A and B, respectively. We find large and consistently positive (albeit noisy) average estimates for large protests and protests on pleasant days, suggesting that large anti-refugee protests inspire counter-mobilization.

In columns 5 and 6, we distinguish between states that belonged to the former GDR (East sample) and those that belonged to the FRG (West sample). Overall, xenophobic sentiments are most pronounced in Eastern Germany – a pattern especially apparent for the right-wing vote (Jaschke et al., 2022). This is in line with the literature connecting a history of socialism with right-wing attitudes (Acemoglu et al., 2022; Lange, 2021). Both TWFE and reducef form estimates are larger in the West than in the East. Given the stronger anti-refugee sentiment in Eastern Germany, a possible explanation is that in regions with pervasive anti-refugee attitudes, do not have the critical mass to mount a counter-movement. In fact, the effect is negative for Eastern Germany in the quasi experimental setting, suggesting that pro-refugee are expressed less on social media potentially because the social cost of pro-refugee activism becomes prohibitively large. Conversely, municipalities in Western Germany experience a rise in pro-refugee activism online in response to protest on pleasant days. Overall, this suggests that large protests can polarize and streamline political action, depending on the relative size of pro- and anti-refugee groups.

4.3 Dynamics of belief updating

Past signals and hate crimes. Traditional models of revolution and their more recent applications to protest, conceptualize the primary cause of revolt as an unexpected large shock, typically in the form of a public signal about a surge in grievances within the population (Davies, 1962, 1978; Correa et al., 2021). The so called J-curve theory suggests that people on aggregate behave like a boiling frog, i.e. if the same condition emerges through gradual change people are less likely to revolt than in the presence of a large shock. This stands in contrast with other models of coordination where beliefs are updated in the same way, irrespective of the frequency and size of the shock (De Mesquita, 2010; Edmond, 2013). In our context, individuals may interpret protest participation as a signal about the overall grievances in the population and decide to revolt by committing hate crimes. We test whether radical action behaves like a boiling frog.

Our baseline specification includes the protest participation in the previous period as a control, capturing the effect of more recent signals on radical action. In Table 4, we investigate the role of past signals more carefully. In column 1, we report the correlation between protest participation and hate crimes from our baseline specification. In columns 2 to 4 we interact protest participation in t with aggregate signals about the grievances of others and control for the interacting variable separately. Column 2 reports the coefficient for the interaction with the cumulative number of past PEGIDA protests. The effect of large protests is magnified by past protest signals. This extends to column 3 where we look at the cumulative (log of one plus) number of previous protest participants. In a last step, we investigate a more direct measure of other people's preferences for radical action in the form of the past cumulative number of hate crimes in the municipality and find that these further encourage hate crimes when public support for the movement is high.

In column 5 and 6, we combine the past signals into one indicator using their first principle component. In column 6, we estimate the reduced form regression based on specification 3 and add the interaction between the past signal index and protest during pleasant weather.[12] The quasi experimental variation in public perception of protests confirm our findings from the previous columns. Overall, we do not find evidence that radical action follows a boiling frog model. On the contrary, municipalities with higher past exposure to public signals intensify radical action, pointing to a continuous update the preferences of others.

Initial protests can catalyze a series of hate crimes, thereby setting off a self-perpetuating cycle of radical action. This cycle is characterized by the fact that hate crimes following one protest not only reflect immediate responses but also serve as a precursor, influencing the likelihood and intensity of future hate crimes. These findings highlight a dynamic interaction where past hate crimes, by reinforcing perceptions of widespread grievances, can escalate and sustain a cycle of violence, perpetuating a feedback loop where each incident fuels the potential for subsequent radical actions.

Persistence of hate crimes. Next, we investigate the timing of hate crimes against refugees. It is possible that extremists do not make a permanent update about the costs and returns to committing hate crimes but have a short-term emotional response to the public support for their cause. In our baseline estimation, we already exclude hate crimes on Mondays to rule out the possibility that large protests animate extremists. In the right panel of Figure 6, we look at the effect of Monday PEGIDA protests during pleasant weather on hate crimes committed on each day of the week. We observe a surge in hate-crimes in the day following the protest but it also persists throughout the weak and peaks on the days leading up to the weekend with the largest effect on Fridays. This is in line with the literature on date-spatial crime predictions that finds a higher crime frequency between Fridays and Sundays (Cohn & Rotton, 2003).

In addition, we investigate the effect on hate crimes in the weeks following the protest in Table A.2. Again, we present TWFE and reduced form estimates for week 1 to 5 following the PEGIDA Monday protest and find that hate crimes persist over time. This confirms our findings from the event study and the previous section, which revealed equally persistent effects of the first PEGIDA protest and social media proximity to large protest locations. Overall, the sustained increase in hate crimes suggests a more strategic choice by perpetrators, rather than a mere immediate emotional reaction to large protests.

[12] We control for pleasant weather, protest, as well as the interaction between pleasant weather and protest with the past signal index.

4.4 Mediating and activating forces

Use of social media. In a next step, we delve into the role of social media. This analysis is presented in Table 5, where we look at three measures of social media usage. The interplay of social media's influence on the impact of far-right protests is twofold, encompassing both supply and demand perspectives. On the supply side, municipalities with higher social media engagement may see more extensive coverage of protests, potentially due to a larger presence of media-savvy individuals or groups that promote these events online. On the demand side, a populace with greater social media consumption may be more exposed to, and thus influenced by, this coverage.

We assess three distinct indicators of social media engagement: the baseline number of Tweets per capita (pooling 2013 and 2014); the number of per capita followers of the AfD's official Facebook account; and the per capita volume of Tweets mentioning PEGIDA in the six months leading up to each protest. These variables are synthesized into a singular index using a first principal component analysis, with the resultant composite measure examined in column 1. Subsequent columns (2 to 4) explore the individual effects of these social media variables, and column 5 conducts a horse race among them.

The results confirm this dual-effect hypothesis: the impact of protest participation is more pronounced in regions with higher social media usage, as indicated by the combined Social Media index and baseline Twitter penetration. The number of Facebook followers of the AfD party, which we employ as a proxy for anti-refugee sentiment, amplifies the effect of protests. This suggests that where there is a higher baseline demand for far-right content, the response to protests is stronger. Conversely, the count of tweets mentioning PEGIDA, which may not unambiguously signal support for the movement, does not significantly impact our findings and shows a negative sign when competing variables are considered. This could reflect a balance of supply-side amplification and demand-side consumption, with the level of active engagement on social media platforms shaping the local response to protests. In sum, our analysis indicates that the effect of far-right protests on hate crimes is amplified in areas where the supply of protest coverage meets a receptive demand among social media users.

Economic grievances. In Panel B of Table 5, we shift our focus to economic grievances as a potential moderator of the effect of protest participation on hate crimes. We employ various proxies to capture the economic strains within municipalities that could potentially galvanize far-right protest participants into committing hate crimes. As in Panel A, the regression analysis incorporates the log-transformed number of protest participants, interacted with the first principal component (PC1) of Economic Grievances, and separately with the share of refugees, the local unemployment rate, and the share of unskilled labor. The coefficients for the interaction between protest participation and the PC1 Economic Grievances, as well as with the refugee share, are statistically significant, suggesting these factors do influence hate crime occurrence. However, the interactions with the unemployment rate and the share of unskilled labor do not demonstrate a significant correlation.

These findings suggest that while the presence of refugees in a municipality is a salient driver of hate crime occurrences following protests, other economic conditions, such as the unemployment rate or the proportion of unskilled labor, do not appear to activate additional grievances that translate into hate crimes. This indicates that the mere presence of refugees might be a sufficient condition to trigger hate crimes in the context of far-right protests, overshadowing other local economic factors. This is in line with recent literature that emphasizes cultural over economic factors in shaping anti-immigrant sentiment (Alesina & Tabellini, 2022).

In Appendix Table A.3, we examine the characteristics of the immigrant population, particularly their economic integration, as indicated by the unemployment ratio between foreigners and natives and the average skill level of foreigners. Again, we include the share of refugees in the set of variables and combine all three into a first principle component index. The results confirm our previous finding that it is the relative number of refugees, rather than the economic assimilation or skill set of immigrants more generally, that predicts hate crimes.

5 Conclusion

In this paper, we examine the interplay between a public signal about the preferences of others – in the form of protest attendance and favorable coverage of the protest – and radical action to advance the same political goal. We investigate this question in the context of right-wing PEGIDA protests and hate crime in Germany allows for the study of political expression and extremist behavior. Our analysis uses a difference-in-differences approach, combined with local weather variations to determine causal effects. We find that protest success and hate crimes act as strategic complements.

While it is very difficult to disentangle the underlying drivers of this complementary, we present several novel pieces of evidence that delineate a reduction in social cost from a change in preferences, i.e. the radicalization of new individuals. First, we document geographic diffusion patterns, highlighting that protest participation across locations do not act as complements in our context. However, protest functions as a signal to individuals that are already sympathetic to the cause, leading to a diffusion of hate crimes through right-wing social media networks. Our results on the shift towards overt and brazen forms of violence as well as our findings on the self-reinforcing mechanisms of past signals and radical action emphasize how even one off protest success can lead to a spiral of radicalization and violence.

Our research adds to the understanding of how non-institutional political activities affect social behavior and norms, particularly in terms of extremist actions within the far-right spectrum. It also emphasizes the significant role of social media in disseminating extremist views and behaviors. Our findings indicate that more moderate protest participants may underestimate the negative externalities of their protest participation for minority groups.

References

Acemoglu, D., De Feo, G., De Luca, G., & Russo, G. (2022). War, Socialism, and the Rise of Fascism: an Empirical Exploration. *The Quarterly Journal of Economics*, 137(2), 1233–1296.

Ajzenman, N., Cavalcanti, T., & Da Mata, D. (2023). More than words: Leaders' speech and risky behavior during a pandemic. *American Economic Journal: Economic Policy*, 15(3), 351–371.

Alesina, A. & Tabellini, M. (2022). *The political effects of immigration: Culture or economics?* Technical report, National Bureau of Economic Research.

Battaglini, M., Morton, R., & Patacchini, E. (2021). Social groups and the effectiveness of petitions.". *mimeo*.

Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of political economy*, 76(2), 169–217.

Berntzen, L. E. & Weisskircher, M. (2016). Anti-islamic pegida beyond germany: Explaining differences in mobilisation. *Journal of Intercultural Studies*, 37(6), 556–573.

Blakeslee, D., Chaurey, R., Fishman, R., Malghan, D., & Malik, S. (2021). In the heat of the moment: Economic and non-economic drivers of the weather-crime relationship. *Journal of Economic Behavior & Organization*, 192, 832–856.

Bray, K., Braakmann, N., & Wildman, J. (2022). Austerity, welfare cuts and hate crime: Evidence from the uk's age of austerity. *Journal of Urban Economics*, (pp. 103439).

Bursztyn, L., Cantoni, D., Yang, D. Y., Yuchtman, N., & Zhang, Y. J. (2021). Persistent political engagement: Social interactions and the dynamics of protest movements. *American Economic Review: Insights*, 3(2), 233–50.

Bursztyn, L., Egorov, G., & Fiorin, S. (2020). From extreme to mainstream: The erosion of social norms. *American economic review*, 110(11), 3522–3548.

Bursztyn, L. & Yang, D. Y. (2022). Misperceptions about others. *Annual Review of Economics*, 14, 425–452.

Cantoni, D., Yang, D. Y., Yuchtman, N., & Zhang, Y. J. (2019). Protests as strategic games: experimental evidence from hong kong's antiauthoritarian movement. *The Quarterly Journal of Economics*, 134(2), 1021–1077.

Carr, J., Clifton-Sprigg, J., James, J., & Vujic, S. (2020). *Love thy neighbour? Brexit and hate crime*. Technical report, IZA Discussion Papers.

Casanueva, A. (2021). Can chants in the street change parliament's tune? the effects of the 15m social movement on spanish elections. *The Effects of the 15M Social Movement on Spanish Elections (April 22, 2021)*.

Chersich, M. F., Swift, C. P., Edelstein, I., Breetzke, G., Scorgie, F., Schutte, F., & Wright, C. Y. (2019). Violence in hot weather: Will climate change exacerbate rates of violence in south africa? *South African Medical Journal*, 109(7), 447–449.

Cohn, E. & Rotton, J. (2003). Even criminals take a holiday: Instrumental and expressive crimes on major and minor holidays. *Journal of Criminal Justice*, 31, 351–360.

Conley, T. G. (1999). Gmm estimation with cross sectional dependence. *Journal of econometrics*, 92(1), 1–45.

Correa, S., Nandong, G., & Shadmehr, M. (2021). Crises, catharses, and boiling frogs: Path dependence in collective action. *Available at SSRN 3906282*.

Davies, J. C. (1962). Toward a theory of revolution. *American sociological review*, (pp. 5–19).

Davies, J. C. (1978). The j-curve theory. *American Political Science Review*, 72(4), 1357–1358.

De Chaisemartin, C. & d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9), 2964–96.

De Chaisemartin, C. & D'Haultfoeuille, X. (2022a). *Difference-in-differences estimators of intertemporal treatment effects*. Technical report, National Bureau of Economic Research.

De Chaisemartin, C. & D'Haultfoeuille, X. (2022b). *Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: A survey*. Technical report, National Bureau of Economic Research.

De Mesquita, E. B. (2010). Regime change and revolutionary entrepreneurs. *American Political Science Review*, 104(3), 446–466.

Dipoppa, G., Grossman, G., & Zonszein, S. (2023). Locked down, lashing out: Covid-19 effects on asian hate crimes in italy. *The Journal of Politics*, 85(2), 389–404.

Edmond, C. (2013). Information manipulation, coordination, and regime change. *Review of Economic studies*, 80(4), 1422–1458.

Entorf, H. & Lange, M. (2019). Refugees welcome? understanding the regional heterogeneity of anti-foreigner hate crimes in germany. *Understanding the Regional Heterogeneity of Anti-Foreigner Hate Crimes in Germany (January 30, 2019). ZEW-Centre for European Economic Research Discussion Paper*, (19-005).

Fages, D. M. & Martínez, A. A. (2023). Hate speech and social media: evidence from bolsonaro's election in brazil.

Field, S. (1992). The effect of temperature on crime. *The British journal of criminology*, 32(3), 340–351.

Goetzmann, W. N., Kim, D., Kumar, A., & Wang, Q. (2015). Weather-induced mood, institutional investors, and stock returns. *The Review of Financial Studies*, 28(1), 73–111.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2), 254–277.

Guhr, O., Schumann, A.-K., Bahrmann, F., & Böhme, H. J. (2020). Training a broad-coverage German sentiment classification model for dialog systems. In N. Calzolari, F. Béchet, P. Blache, K. Choukri, C. Cieri, T. Declerck, S. Goggi, H. Isahara, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, & S. Piperidis (Eds.), *Proceedings of the Twelfth Language Resources and Evaluation Conference* (pp. 1627–1632). Marseille, France: European Language Resources Association.

Guriev, S. & Papaioannou, E. (2022). The political economy of populism. *Journal of Economic Literature*, 60(3), 753–832.

Han, S., Riddell, J. R., & Piquero, A. R. (2023). Anti-asian american hate crimes spike during the early stages of the covid-19 pandemic. *Journal of Interpersonal Violence*, 38(3-4), 3513–3533. PMID: 35657278.

Hanes, E. & Machin, S. (2014). Hate crime in the wake of terror attacks: Evidence from 7/7 and 9/11. *Journal of Contemporary Criminal Justice*, 30(3), 247–267.

Heilmann, K., Kahn, M. E., & Tang, C. K. (2021). The urban crime and heat gradient in high and low poverty areas. *Journal of Public Economics*, 197, 104408.

Jaschke, P., Sardoschau, S., & Tabellini, M. (2022). *Scared Straight? Threat and Assimilation of Refugees in Germany*. Technical report, National Bureau of Economic Research.

Jiang, J., Murrugara-Llerena, N., Bos, M. W., Liu, Y., Shah, N., Neves, L., & Barbieri, F. (2022). Sunshine with a chance of smiles: How does weather impact sentiment on social media? In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 16 (pp. 393–404).

Kanol, E. & Knoesel, J. (2021). Right-wing extremist mobilization in germany. *WZB Berlin Social Science Center, Version*, 1(0), 125–41.

Lagios, N., Méon, P.-G., & Tojerow, I. (2022). Is demonstrating against the far right worth it? evidence from french presidential elections.

Lange, M. (2021). The Legacy of State Socialism on Attitudes Toward Immigration. *Journal of Comparative Economics*, 49(3), 733–750.

Larreboure, M. & González, F. (2021). The impact of the women's march on the us house election. *Pontificia Universidad Catolica de Chile,(Maret 2021)*, (pp. 4–5).

Levy, R. & Mattsson, M. (2023). The effects of social movements: Evidence from# metoo. *Available at SSRN 3496903*.

Madestam, A., Shoag, D., Veuger, S., & Yanagizawa-Drott, D. (2013). Do political protests matter? evidence from the tea party movement. *The Quarterly Journal of Economics*, 128(4), 1633–1685.

Mellon, J. (2021). Rain, rain, go away: 176 potential exclusion-restriction violations for studies using weather as an instrumental variable. *Available at SSRN 3715610*.

Michel, S. J., Wang, H., Selvarajah, S., Canner, J. K., Murrill, M., Chi, A., Efron, D. T., & Schneider, E. B. (2016). Investigating the relationship between weather and violence in baltimore, maryland, usa. *Injury*, 47(1), 272–276.

Miller, D. L. (2023). An introductory guide to event study models. *Journal of Economic Perspectives*, 37(2), 203–230.

Müller, K. & Schwarz, C. (2021). Fanning the flames of hate: Social media and hate crime. *Journal of the European Economic Association*, 19(4), 2131–2167.

Müller, K. & Schwarz, C. (2023). From hashtag to hate crime: Twitter and antiminority sentiment. *American Economic Journal: Applied Economics*, 15(3), 270–312.

OECD (2021). *Perspectives on Global Development 2021*.

Ortoleva, P. & Snowberg, E. (2015). Overconfidence in political behavior. *American Economic Review*, 105(2), 504–535.

Passarelli, F. & Tabellini, G. (2017). Emotions and political unrest. *Journal of Political Economy*, 125(3), 903–946.

Potrafke, N. & Roesel, F. (2022). Online versus offline: Which networks spur protests?

Qin, B., Strömberg, D., & Wu, Y. (2021). Social media and collective action in china.

Romarri, A. (2020). Do far-right mayors increase the probability of hate crimes? evidence from italy. *Evidence From Italy (December 6, 2020)*.

Roth, J., Sant'Anna, P. H., Bilinski, A., & Poe, J. (2022). What's trending in difference-in-differences? a synthesis of the recent econometrics literature. *arXiv preprint arXiv:2201.01194*.

Sangnier, M. & Zylberberg, Y. (2017). Protests and trust in the state: Evidence from african countries. *Journal of Public Economics*, 152, 55–67.

Sarsons, H. (2015). Rainfall and conflict: A cautionary tale. *Journal of development Economics*, 115, 62–72.

Sharoff, S. (2006). Open-source corpora: Using the net to fish for linguistic data. *International Journal of Corpus Linguistics*, 11(4), 435–462.

Vorländer, H., Herold, M., & Schäller, S. (2018). *PEGIDA and new right-wing populism in Germany*. Springer.

Wasow, O. (2020). Agenda seeding: How 1960s black protests moved elites, public opinion and voting. *American Political Science Review*, 114(3), 638–659.

Wooldridge, J. (2021). Two-way fixed effects, the two-way mundlak regression, and difference-in-differences estimators. *Available at SSRN 3906345*.

Xu, R., Xiong, X., Abramson, M. J., Li, S., & Guo, Y. (2020). Ambient temperature and intentional homicide: A multi-city case-crossover study in the us. *Environment international*, 143, 105992.

Zhong, B. & Zhou, Y. (2012). "under the weather": The weather effects on us newspaper coverage of the 2008 beijing olympics. *Mass Communication and Society*, 15(4), 559–577.

6 Tables and Figures

Figure 1. Cumulative number of protest and hate-crimes over time

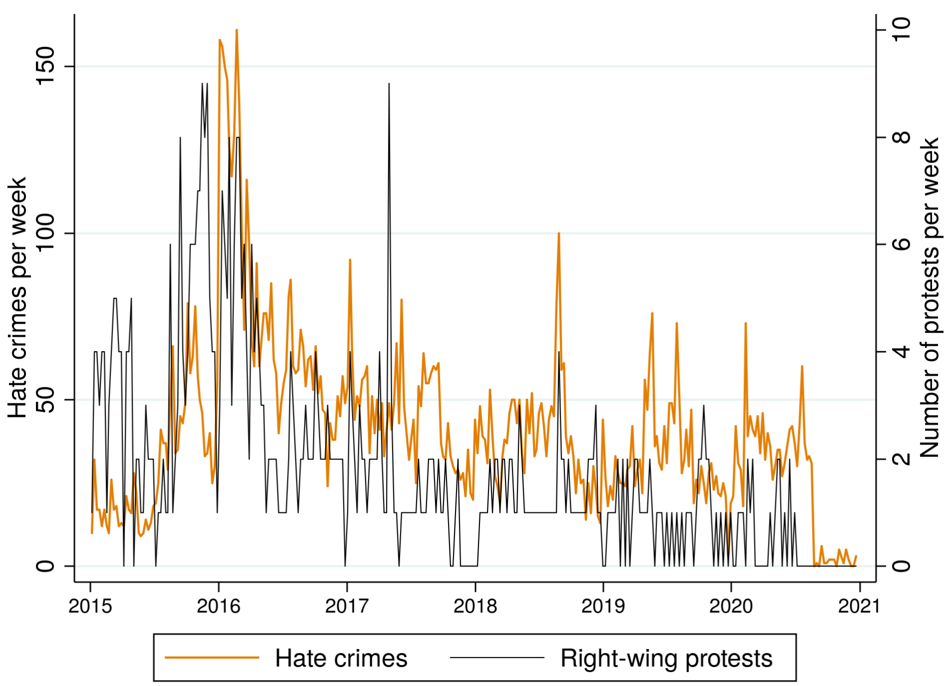


Figure 2. Cumulative number of PEGIDA Monday protest participants and hate crimes per 100K (2015-2020)

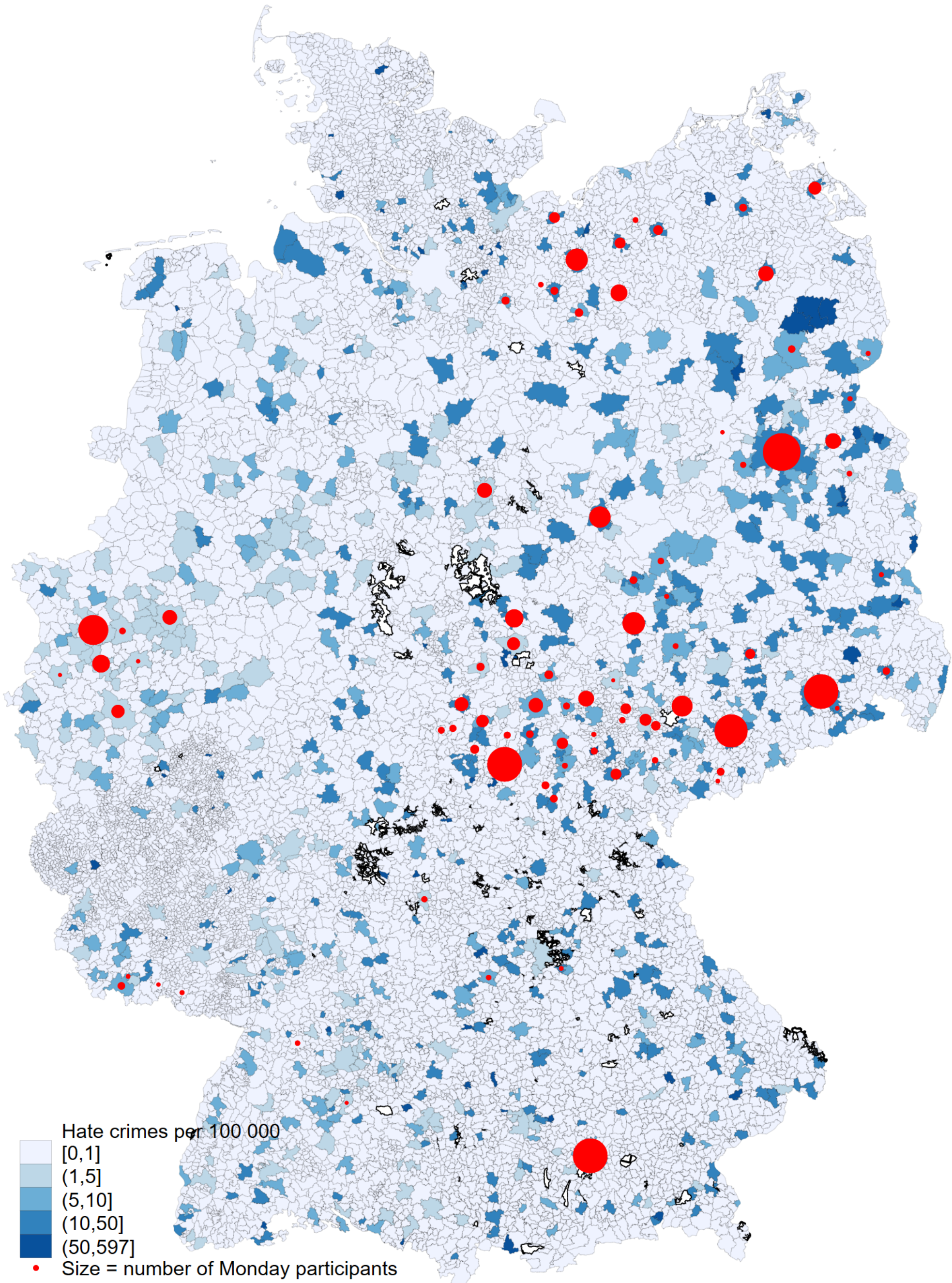
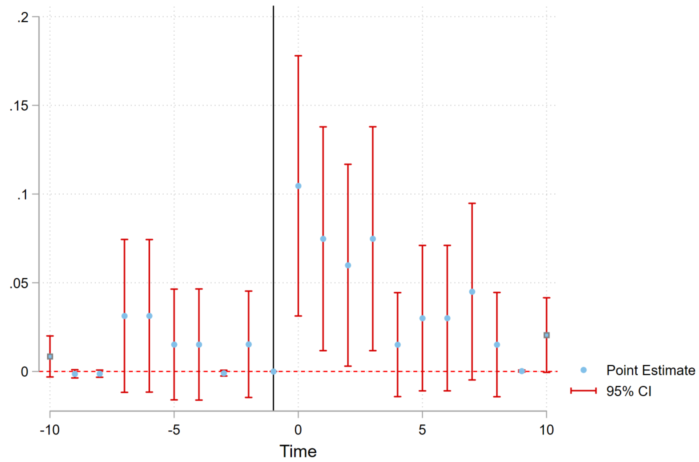
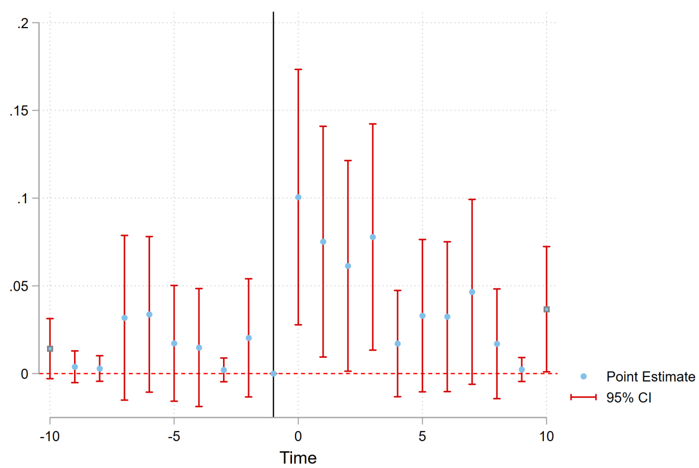


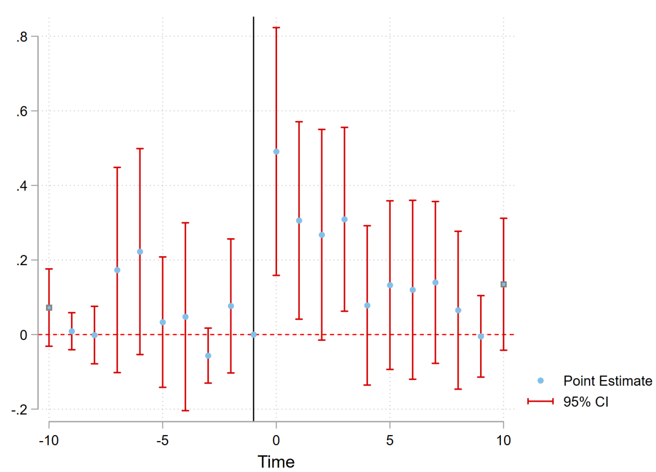
Figure 3. Event Study: Timing of first PEGIDA Monday Protest and Hate Crimes against Refugees



(a) Full sample of municipalities (never treated and later treated sample)

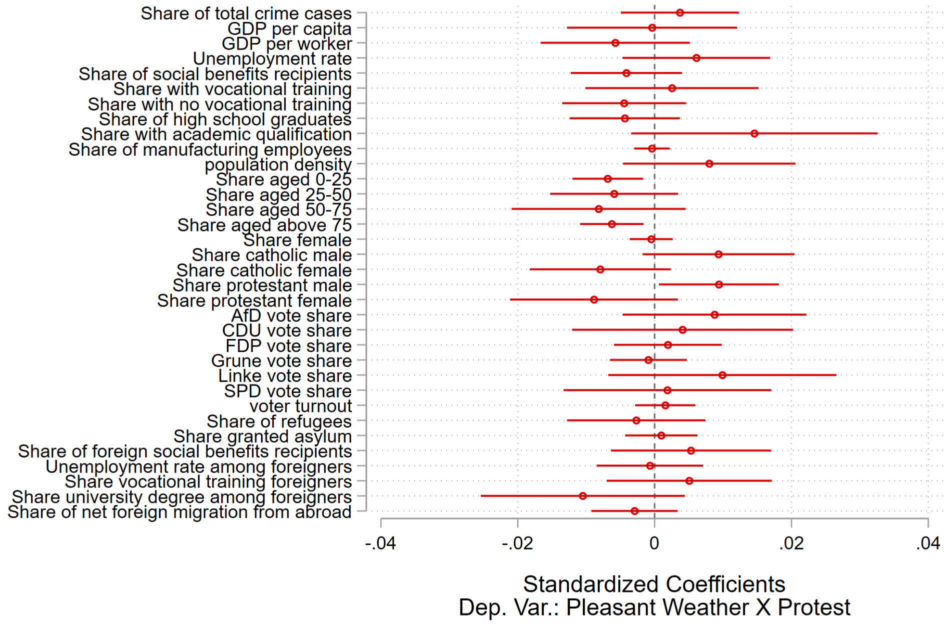


(b) Sample of municipalities with protests (ever treated sample)



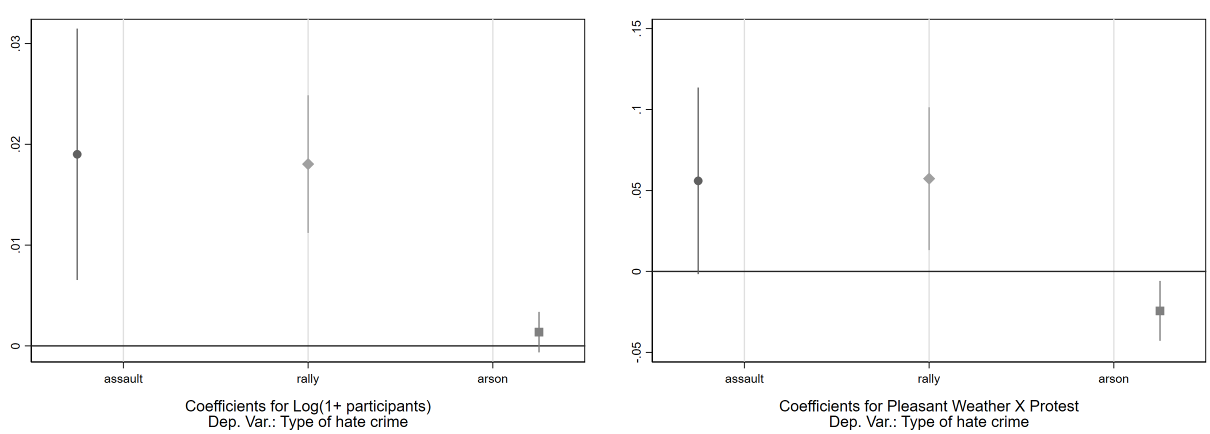
(c) Sample of municipalities with hate crimes

Figure 4. Municipality characteristics and protest on pleasant days



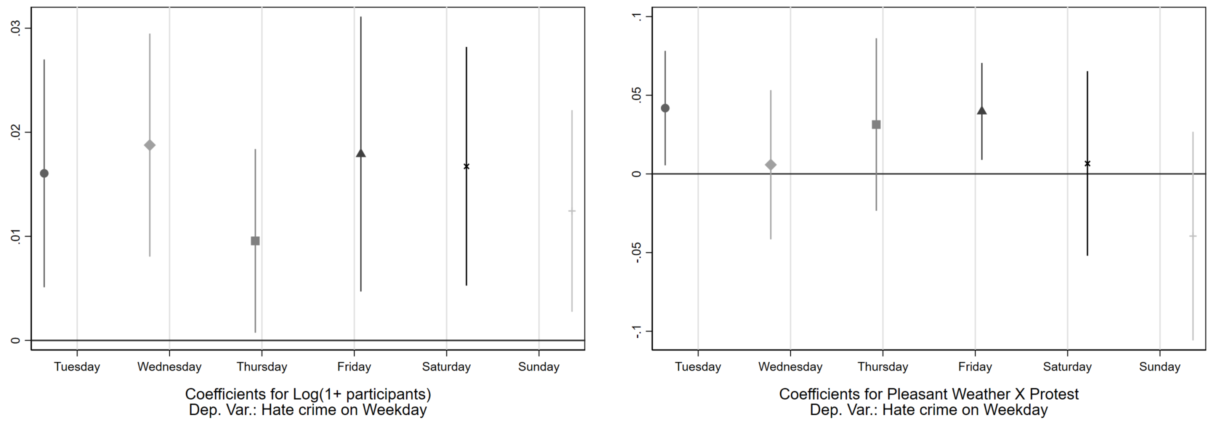
Note: Plausibility of the instrument exogeneity. We estimate the following regression: $\sum Pleasant \times Protest_{muni.year} = \beta X_{muni.year} + \sum Pleasant_{muni.year} + \sum Protest_{muni.year} + \gamma_{state} + \theta_{year} + \epsilon_{muni.year}$. We regress the set of baseline control variables and additional municipality characteristics on the cumulative number of protests on pleasant days in a municipality and year. We include state fixed effects as well as year fixed effects and control for the cumulative number of pleasant days and cumulative number of Monday protests separately. Standard errors are clustered at the state level. Coefficients are standardized for comparability. We re-scale the following variables in a 1:2 ratio for readability: share with academic qualifications, share of foreigners with academic qualifications, share of protestants and catholics (both male and female).

Figure 5. Type of Hate Crimes: shift to overt and brazen forms of violence



Note: Dependent variable: type of hate crime committed in the 6 days following the Monday protest.

Figure 6. Timing of Hate Crimes



Note: Coefficient plot for protest on a pleasant day. We estimate the following regression (same as baseline): $Hate\ Crime_{it}^{timing/type} = \beta Protest_{it} \times Weather_{it} + \eta_1 Protest_{it} \times X'_{it} + \eta_2 Weather_{it} \times X'_{it} + \mu_{im} + \delta_t + \epsilon_{it}$ We include week and municipality month-fixed effects, full set of interacted control variables. Dependent variable: any hate crime committed on a specific weekday (excluding Monday, i.e. protest day itself).

Table 1. **Summary statistics**

	Mean	Sd	Min	Max
Main Variables				
Hate crime (probability)	0.0002	0.0130	0	1
Protest (probability)	0.0002	0.0142	0	1
Total participants	0.0291	5.1182	0	6000
Log(1 + participants)	0.0009	0.0649	0	8.6997
Total participants — protest	145.42	331.50	4	6000
Instrument				
Pleasant weather × Protest	0.0001	0.0119	0	1
Pleasant weather (probability)	0.6626	0.4728	0	1
Controls				
GDP per capita	3.15e+04	9186.5310	1.51e+04	1.88e+05
Population density	217.7027	311.8760	3.4665	4736.1055
Unemployment rate	0.0201	0.0111	0.0000	0.2430
Share of refugees	0.0110	0.0061	0.0004	0.1300
AfD vote share	0.0975	0.0673	0.0039	0.5000
Share of total crime cases	0.0447	0.0286	0.0130	1.0398
Share with no vocational training	0.0353	0.0141	0.0064	0.1427
Share aged 0-25	0.2377	0.0326	0.0969	0.5288
Share aged 25-50	0.3044	0.0272	0.1397	0.4675
Share aged 50-75	0.3488	0.0358	0.0825	0.6000
Share female	0.5032	0.0163	0.0859	0.6356
Unemployment rate among foreigners	0.0027	0.0029	0.0000	0.0538
Share university degree among foreigners	0.0027	0.0030	0.0002	0.0441
Share granted asylum	0.0059	0.0040	0.0000	0.1024
#RefugeesWelcome tweets per capita (baseline)	0.0000	0.0000	0.0000	0.0004
Tweet frequency per capita (baseline)	0.3555	0.3153	0.0042	3.4582
AfD followers (baseline)	0.0001	0.0001	0.0000	0.0033
Interacting variables				
Tweets PEGIDA (per capita, last 6 mo.)	0.1526	0.3478	0.0000	16.0755
Observations	2,156,734			

Note: Descriptive statistics (mean, standard deviation, min and max) of all variables used for the analysis divided by the function they play in the empirical strategy.

Table 2. **Protest Participation increases probability of hate crimes**

	baseline controls (1)	+ demographic controls (2)	+ refugee controls (3)	+ social media controls (4)
Panel A: Protest Participation				
Log(1 + participants)	0.0626***	0.0626***	0.0626***	0.0630***
	(0.0163)	(0.0163)	(0.0163)	(0.0164)
Adj. R-squared	0.309	0.309	0.309	0.310
Municipality FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes
Baseline controls	Yes	Yes	Yes	Yes
Demographic controls		Yes	Yes	Yes
Refugee controls			Yes	Yes
Social Media controls				Yes
Panel B: Protest on pleasant days				
Weather × Protest	0.0779***	0.0667***	0.0665***	0.0687***
	(0.0218)	(0.0200)	(0.0199)	(0.0203)
Weather	6.79e-06	0.000664	0.000599	0.000638
	(9.17e-05)	(0.000546)	(0.000543)	(0.000588)
Protest	0.577***	1.495	3.331	4.274
	(0.134)	(4.053)	(4.076)	(4.263)
Adj. R-squared	0.377	0.380	0.381	0.383
Municipality × month of year FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes
Weather & Protest × baseline controls	Yes	Yes	Yes	Yes
Weather & Protest × demographic controls		Yes	Yes	Yes
Weather & Protest × refugee controls			Yes	Yes
Weather & Protest × social media controls				Yes
Observations	2,290,440	2,290,440	2,290,440	2,156,734
Municipalities	9,688	9,688	9,688	9,023
Mean dep. var.	0.000168			

Note: SE clustered by municipality; *p<0.1; **p<0.05; ***p<0.01. Time horizon is January 2015 until December 2019. Panel A: two-way fixed effects regression with municipality and week fixed effects. Main treatment is measured as the log of 1 + total participants at a PEGIDA Monday demonstration using data from Kanol & Knoesel (2021). Panel B: two-way fixed effects regression with municipality-month of the year and week fixed effects. Main treatment is the interaction between pleasant weather conditions during a Monday and the presence of protest in that week. In this panel, we additionally include weather and protest separately. We also include the interaction of all the controls with the both weather conditions and the presence of protest. For both panels; outcome comes from ProAsyl and AAF and is measured as a dummy variable equals to 1 if any hate crime against refugees was committed between Tuesday and Sunday of the same week. Municipality-level base controls include: the lagged log (1 +) number of protest participants in the previous week and the lagged dependent variable, GDP per capita; population density; unemployment share; refugee share, latest national vote shares for the AfD and crime rate per 100K inhabitants. Demographic controls include: share of unskilled; share of females; dummies for share of population aged 0-25 25-50 and 50-75. Refugee controls include: share of foreign unemployed; shared of skilled foreigners; share of asylum status granted. Social media controls include: the baseline Twitter penetration at the NUTS-3 level in 2014, the number of tweets containing #refugeesWelcome at the NUTS-3 level per capital in 2014 and the number of Facebook users following AfD before 2015 scaled by population (taken from Müller & Schwarz, 2021) at the groups of municipality level.

Table 3. **Diffusion of public signal: geographic and (right-wing) social media proximity to other large protest locations**

Outcome:	Protest [0,1]	Log (1 + participants)		Hate Crime [0,1]	
Timing outcome:	t	t	t + 1	t	t + 1
	(1)	(2)	(3)	(4)	(5)
Panel A: time varying social media network - retweets in the previous 6 months					
S_{it}^{P} (PEGIDA social media proximity)	0.000359*	0.00194	0.00178	0.000568***	0.000583***
	(0.000199)	(0.00123)	(0.00118)	(0.000178)	(0.000182)
S_{it}^{W} (wider social media proximity)	7.72e-05	0.000495	0.000635	5.57e-05	8.67e-05
	(0.000226)	(0.000934)	(0.00106)	(6.79e-05)	(8.76e-05)
G_{it} (geographic proximity)	-2.78e-05	-0.000125	-1.52e-05	-5.46e-07	-1.81e-05
	(2.99e-05)	(0.000143)	(0.000131)	(1.61e-05)	(1.40e-05)
$Log(1 + participants)_{it}$				0.0521***	0.0274***
				(0.0102)	(0.00943)
Panel B: constant social media network - retweets/followers at baseline					
S_{it}^{P}	5.05e-05	0.000315	0.000518	0.000221*	0.000230*
	(0.000147)	(0.000668)	(0.000692)	(0.000131)	(0.000126)
S_{it}^{W}	3.04e-05	0.000150	0.000204	3.36e-05*	3.67e-05**
	(4.02e-05)	(0.000186)	(0.000198)	(1.74e-05)	(1.87e-05)
G_{it}	-1.96e-05	-7.71e-05	2.70e-05	3.47e-06	-1.24e-05
	(2.49e-05)	(0.000122)	(9.86e-05)	(1.66e-05)	(1.05e-05)
$Log(1 + participants)_{it}$				0.0524***	0.0277***
				(0.0103)	(0.00952)
Observations	2,280,752	2,280,752	2,271,064	2,280,752	2,271,064
Municipalities	9688	9688	9688	9688	9688
Base controls & FE	Yes	Yes	Yes	Yes	Yes

Note: SE clustered by municipality; *p<0.1; **p<0.05; ***p<0.01. Time horizon is January 2015 until December 2019. Two-way fixed effects regression with municipality and week fixed effects. S_it is the social media network proximity-weighted sum of protest participants in other municipalities (standardized). Superscript P indicates PEGIDA network, superscript W indicators wider network, based on random sample of tweets. G_it geographic proximity weighted (linearly decreasing window function until 100km) sum of participants in other municipalities (standardized). All columns include the full set of base controls (see Table 2) including lagged protest participation and lagged hate crimes as independent variables. Outcome in column 1 is any protest on the same day, columns 2 and 3 is the log 1 + number of participants in the same week and the following week respectively. Outcome in columns 4 is any hate crime committed in the same week, and the following week in columns 5. Panel A measures social media proximity as the per capita number of retweets in municipality i over the previous six months of sampled tweets coming from municipality j. Panel B uses retweets at baseline, i.e. before January 2015 for the wider social media proximity and follower relationship at baseline for the right-wing social media proximity.

Table 4. **Dynamic effects: past protest success increases hate crimes**

Outcome:	any hate crime [0,1]					
	(1)	(2)	(3)	(4)	(5)	(6)
Log(1 + participants)$_{it}$	0.0524***	0.0336***	-0.0632	0.0422***	0.0352***	
	(0.0102)	(0.00798)	(0.0390)	(0.00728)	(0.0105)	
Log(1 + participants)$_{it}$ × past cumul. protest$_{it}$		0.00120***				
		(0.000180)				
Log(1 + participants)$_{it}$ × past cumul. participation$_{it}$			0.0167**			
			(0.00655)			
Log(1 + participants)$_{it}$ × past cumul. hate crimes$_{it}$				0.000340***		
				(4.82e-05)		
Log(1 + participants)$_{it}$ × past signal (PC1)$_{it}$					0.00174***	
					(0.000199)	
Weather × Protest						0.0440
						(0.0427)
Weather × Protest × past signal (PC1)$_{it}$						0.0012***
						(0.000580)
Observations	2,280,752	2,280,752	2,280,752	2,280,752	2,280,752	2,290,440
Municipalities	9688	9688	9688	9688	9688	9688
Adj. R-squared	0.325	0.347	0.336	0.351	0.345	0.378
Municipality FE	Yes	Yes	Yes	Yes	Yes	
Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Baseline controls	Yes	Yes	Yes	Yes	Yes	
Interacting var.	Yes	Yes	Yes	Yes	Yes	Yes
Municipality × month of year FE						Yes
Protest, Weather, PC1 × baseline controls						Yes

Note: SE clustered by municipality; *p<0.1; **p<0.05; ***p<0.01. Time horizon is January 2015 until December 2019. Two-way fixed effects regression with municipality and week fixed effects. Main treatment is measured as the log of 1 + total participants at a PEGIDA Monday demonstration using data from Kanol & Knoesel (2021). Outcome comes from ProAsyl and AAF and is measured as a dummy variable equals to 1 if any hate crime against refugees was committed between Tuesday and Sunday of the same week. Municipality-level base controls include: the lagged log (1+) number of protest participants in the previous week and the lagged dependent variable, GDP per capita; population density; unemployment share; refugee share, latest national vote shares for the AfD and crime rate per 100K inhabitants. Interacting variables are the number of cumulative protests in municipality i until week t in column 2, the log 1 + cumulative number of past protest attendees in column 3, and the cumulative number of hate crimes in column 4. We always control for the interacting variable.

Table 5. **Local characteristics and hate crimes: social media penetration and economic grievances**

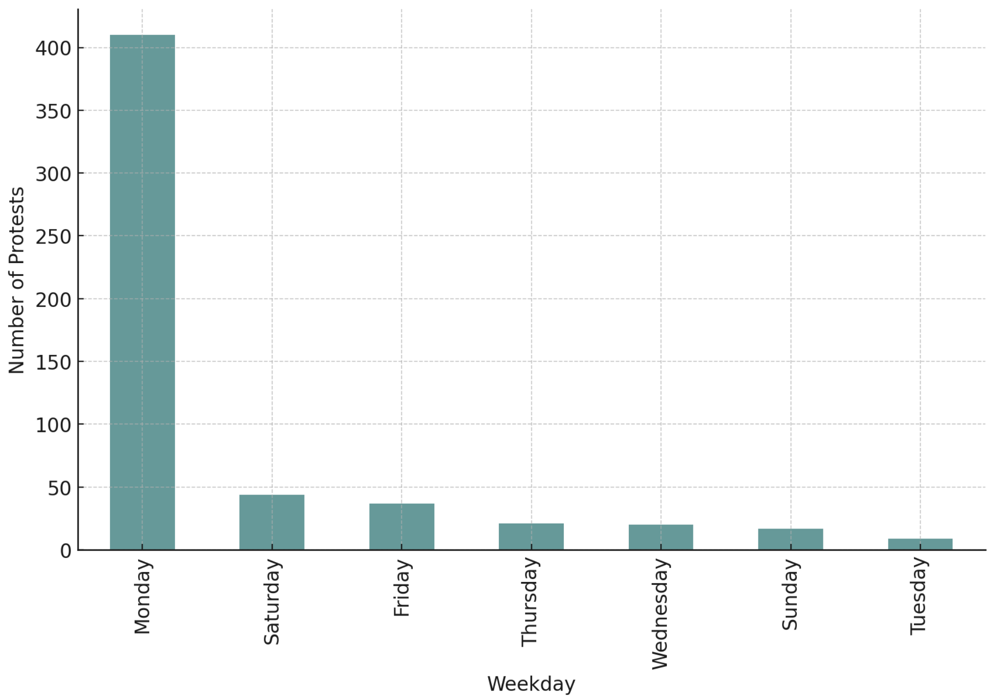
	(1)	(2)	(3)	(4)	(5)
Panel A: Social Media Penetration					
Log(1 + participants)	0.0288***	0.0228***	0.0276***	0.0428***	0.0148**
	(0.00752)	(0.00624)	(0.00920)	(0.00811)	(0.00696)
× PC1 Social Media	0.00711**				
	(0.00283)				
× Tweet penetration (baseline, per capita)		0.0119***			0.0127***
		(0.00199)			(0.00226)
× AfD followers			0.0306**		0.0140*
			(0.0128)		(0.00782)
× Tweets PEGIDA (last 6 months, per capita)				0.00335	-0.00185
				(0.00260)	(0.00229)
Interacting variable	-0.000252			-0.000135	
	(0.000155)			(9.00e-05)	
Adj. R-squared	0.344	0.355	0.332	0.330	0.359
Observations	2,147,711	2,280,752	2,147,711	2,280,752	2,147,711
Municipalities	9,023	9,688	9,023	9,688	9,023
Mean dep. var.	0.000169	0.000160	0.000169	0.000160	0.000169
Panel B: Economic Grievances					
Log(1 + participants)	0.0581***	0.0430***	0.0579***	0.0521***	0.0427***
	(0.0104)	(0.00715)	(0.0155)	(0.0101)	(0.0140)
× PC1 Economic Grievance	0.0125***				
	(0.00433)				
× Refugee share		0.0153***			0.0182***
		(0.00354)			(0.00465)
× Unemployment rate			-0.00236		-0.000478
			(0.00612)		(0.00605)
× Share unskilled				0.00571	-0.00654
				(0.00555)	(0.00554)
Interacting variable	-0.000191	-7.23e-05	1.41e-06	-2.20e-05	
	(0.000502)	(5.82e-05)	(2.43e-05)	(0.000294)	
Adj. R-squared	0.332	0.338	0.325	0.326	0.340
Observations	2,280,752	2,280,752	2,280,752	2,280,752	2,280,752
Municipalities	9,688	9,688	9,688	9,688	9,688
Mean dep. var.	0.000169	0.000169	0.000169	0.000169	0.000169
Municipality FE	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes
Base controls	Yes	Yes	Yes	Yes	Yes

Note: SE clustered by municipality; *p<0.1; **p<0.05; ***p<0.01. Time horizon is January 2015 until December 2019. Two-way fixed effects regression with municipality and week fixed effects, and interaction of the main treatment with characteristics at the municipality-week level. All interacted variables are standardized. In Panel A, column 2 uses Twitter penetration measured at baseline from a random sample of German-language tweets, column 3 the number of AfD followers per capita before 2015 from Müller & Schwarz (2021), and column 4 the per-capita number of tweets mentioning PEGIDA in the last 6 months. In Panel B, column 2 uses the share of refugees in the municipality, column 3 the unemployment rate, and column 4 the share of unskilled workers. In both panels, column 1 uses as interacting variable the first principal components of the three variables, and column 5 presents a horse-race between the three interacting variables. Main treatment is measured as the log of 1 + total participants at a PEGIDA Monday demonstration using data from Kanol & Knoesel (2021). Outcome comes from ProAsyl and AAF and is measured as a dummy variable equals to 1 if any hate crime against refugees was committed between Tuesday and Sunday of the same week. Municipality-level controls include: the lagged log (1 +) number of protest participants in the previous week and the lagged dependent variable; GDP per capita; population density; unemployment share; refugee share; latest national vote shares for the AfD and crime rate per 100K inhabitants.

Appendix

Appendix A: Additional Tables and Figures

Figure A.1. Number of PEGIDA (and offshoot) Protests by Weekday (2015-2020)



Note: Sum over the period 2015-2019 of the number of PEGIDA protest per day of the week.

Figure A.2. Frequency Histogram: Timing of first PEGIDA Monday Protest (by yearly quarters)

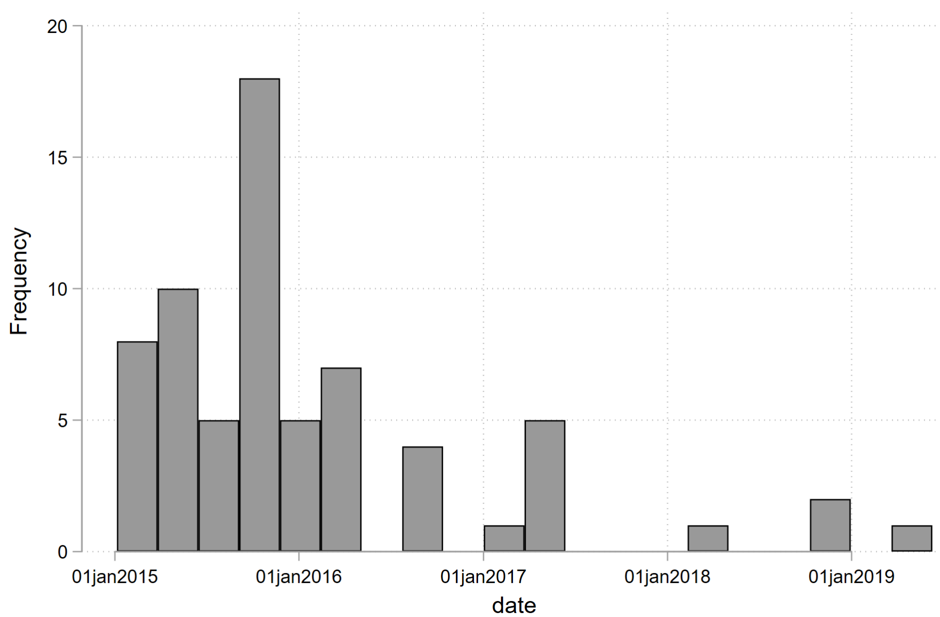
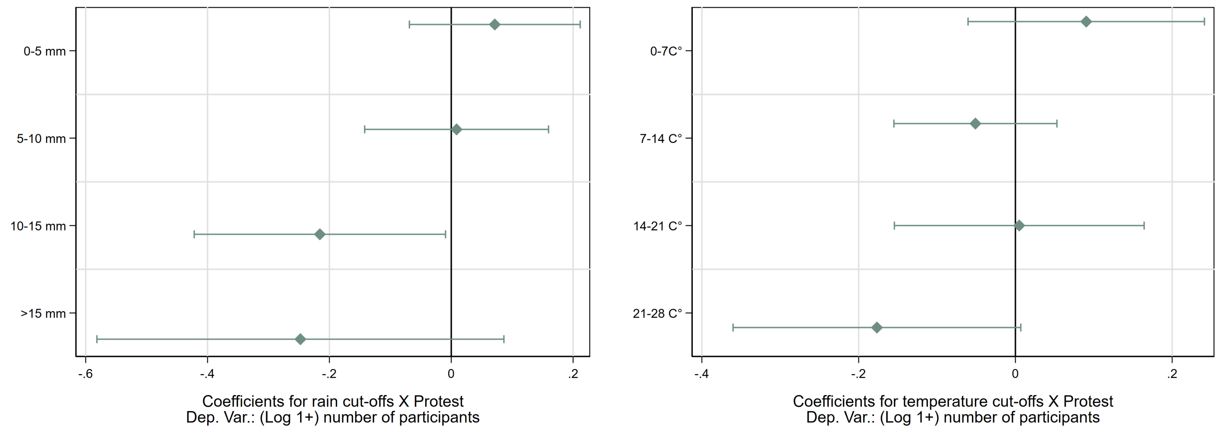
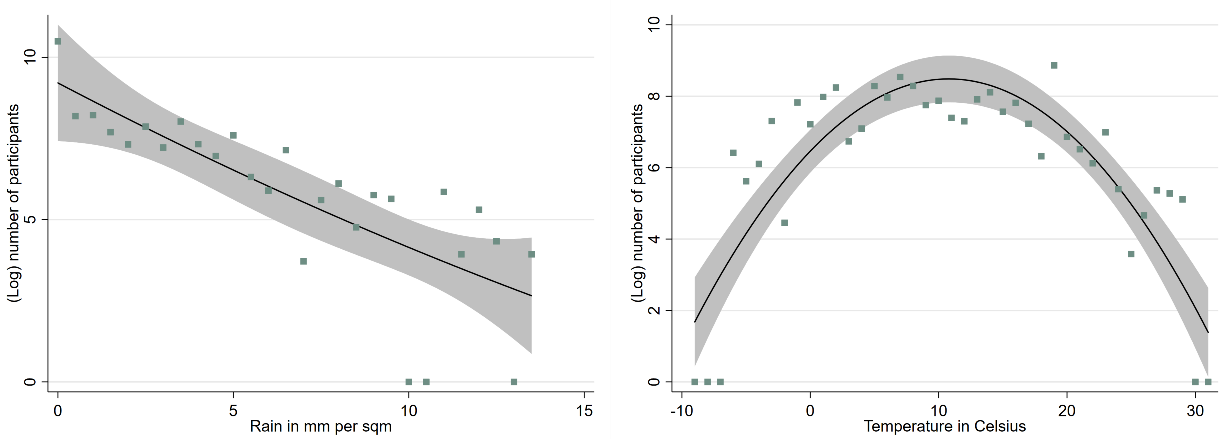


Figure A.4. Coefficient plot: protest participation and weather conditions



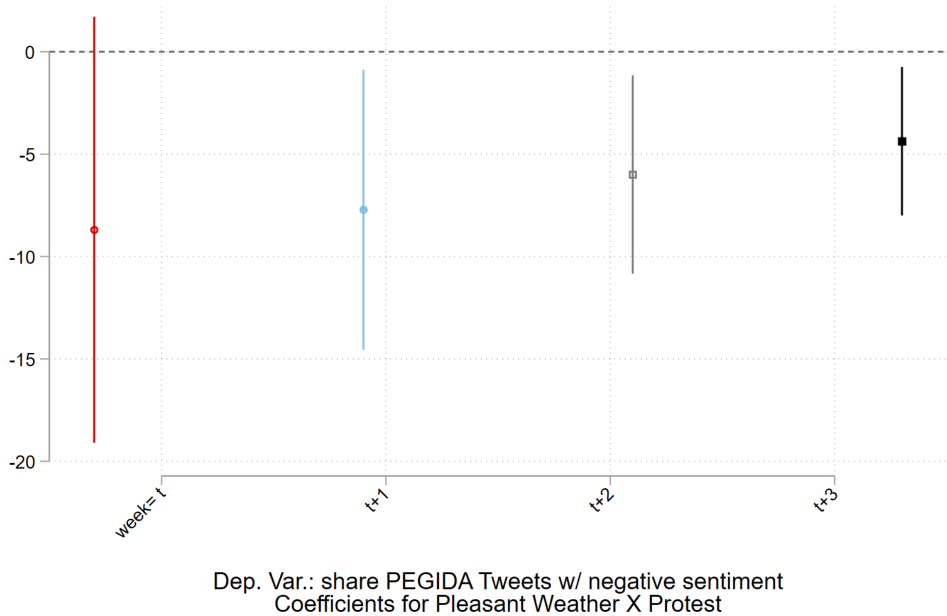
Note: Coefficient plot for the interaction between protest day and various weather cut-offs. We estimate the following regression: $Log(1+participants)_{it} = \beta Protest_{it} \times Weather\ Cut_{it} + \eta_1 Protest_{it} \times X'_{it} + \eta_2 Weather\ Cut_{it} \times X'_{it} + \mu_{im} + \delta_t + \epsilon_{it}$. We estimate separate regressions for each cut-off variable, comparing participation within this weather cut-off to participation within all other cut-offs. Dependent variable is the log of 1+ number of participants. We include municipality month of the year fixed effects μ_{im} as well as week fixed effects δ_t. We also include the weather cut-off variable and the protest variable interacted with the full set of controls ($Weather\ Cut_{it} \times X'_{it}$ and $Protest_{it} \times X'_{it}$). Standard errors are clustered at the municipality level. 95 percent confidence intervals are reported. Left panel uses precipitation cut-offs. We create a dummy variable that switches on if the maximum precipitation between 12pm and 8pm lies between 0-5 mm, 5-10mm, 10-15mm, $>$ 15mm and run separate regressions with each cut-off as the $Weather\ Cut_{it}$ variable. Right panel shows cut-offs for average temperature on the protest day between 12pm and 8pm for 0-5, 5-10, 10-15, 15-20, 20-25 and 25-30 degrees Celsius.

Figure A.3. Scatter plot: protest participation and weather conditions



Note: Left panel shows the unconditional scatterplot and fitted line with 95% confidence intervals between the log of 1 + the number of participants at a Monday PEGIDA protest on the (binned) precipitation measured as an average rain in mm per sqm between 12pm and 8pm on the protest day. The right panel repeats this analysis, this time using the average temperature in Celsius between 12 pm and 8pm on the protest day.

Figure A.5. Protest on Pleasant Days and Negative Sentiment of PEGIDA Tweets



Note: Reduced form regression based on Specification 3. Outcome is measured as the share of tweets at the NUTS3-level with a negative sentiment in the same week as the protest t, as well as in the three weeks following the protest. Sentiment analysis using machine learning tool BERTA based on a German language dictionary.

Table A.1. **Backlash effect: PEGIDA protest increase the number of tweets #RefugeesWelcome**

	Base controls (1)	+ Demographic controls (2)	+ Refugee controls (3)	+ Social media controls (4)	East sample (5)	West sample (6)
Panel A: Protest Participation						
Log(1 + participants)	0.166*	0.165*	0.167*	0.169*	0.194**	0.215
	(0.0866)	(0.0864)	(0.0857)	(0.0861)	(0.0943)	(0.187)
Adj. R-squared	0.199	0.199	0.199	0.200	0.302	0.198
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes
Demographic controls		Yes	Yes	Yes		
Refugee controls			Yes	Yes		
Social media controls				Yes		
Panel B: Protest on pleasant days						
Pleasant weather × Protest	0.720	0.594	0.660	0.650	-1.305*	2.931*
	(0.815)	(0.821)	(0.746)	(0.678)	(0.754)	(1.531)
Weather	-0.0316***	0.131	0.171	0.126	-0.137***	-0.0204**
	(0.00844)	(0.111)	(0.111)	(0.119)	(0.0421)	(0.0101)
Protest	0.750	-194.2**	-98.71	-175.8**	-2.269	-0.334
	(4.012)	(88.81)	(72.84)	(71.06)	(2.888)	(6.963)
Adj. R-squared	0.225	0.225	0.225	0.226	0.339	0.224
Municipality × month of year FE	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Weather & Protest × baseline controls	Yes	Yes	Yes	Yes	Yes	Yes
Weather & Protest × demographic controls		Yes	Yes	Yes		
Weather & Protest × refugee controls			Yes	Yes		
Weather & Protest × social media controls				Yes		
Observations	1,851,143	1,851,143	1,851,143	1,739,629	288,250	1,562,893
Municipalities	9,677	9,677	9,677	9,013	1,619	8,058
Mean dep. var.	0.000160	0.000160	0.000160	0.000160	0.000469	0.000104

Note: SE clustered by municipality; *p<0.1; **p<0.05; ***p<0.01. Panel A: two-way fixed effects regression with municipality and week fixed effects. Main treatment is measured as the log of 1 + total participants at a PEGIDA Monday demonstration using data from Kanol & Knoesel (2021). Panel B: two-way fixed effects regression with municipality-month of the year and week fixed effects. Main treatment is the interaction between pleasant weather conditions during a Monday and the presence of protest in that week. In this panel, we additionally include weather and protest separately. We also include the interaction of all the controls with the both weather conditions and the presence of protest. For both panels: Outcome variable is the number of #RefugeesWelcome tweets per capita measured at the NUTS-3 level and only until the end of 2018. Municipality-level base controls include: the lagged log (1 +) number of protest participants in the previous week and the lagged dependent variable, GDP per capita; population density; unemployment share; refugee share, latest national vote shares for the AfD and crime rate per 100K inhabitants. Demographic controls include: share of unskilled; share of females; dummies for share of population aged 0-25 25-50 and 50-75. Refugee controls include: share of foreign unemployed; shared of skilled foreigners; share of asylum status granted. Social media controls include: the baseline Twitter penetration at the NUTS-3 level in 2014, the number of tweets containing #refugeesWelcome at the NUTS-3 level per capital in 2014 (at baseline) and the number of Facebook users following AfD before 2015 scaled by population (taken from Müller & Schwarz, 2021) at the groups of municipality level.

Table A.2. **The effect of protest on hate crime persist for several weeks**

	Hate crimes x weeks after					
	same week	1	2	3	4	5
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Protest Participation						
Log(1 + participants)	0.0524***	0.0272***	0.0323***	0.0239**	0.0282***	0.0189**
	(0.0102)	(0.00904)	(0.0105)	(0.00993)	(0.0101)	(0.00889)
Adj. R-squared	0.325	0.292	0.287	0.280	0.278	0.273
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Base controls	Yes	Yes	Yes	Yes	Yes	Yes
Panel B: Protest on pleasant days						
Pleasant weather × Protest	0.0793***	0.0779**	0.0431*	0.00972	0.0378	0.0193
	(0.0263)	(0.0394)	(0.0233)	(0.0475)	(0.0342)	(0.0289)
Weather	3.40e-05	-4.83e-05	9.57e-05	-0.000236*	0.000110	0.000197
	(9.73e-05)	(0.000117)	(0.000158)	(0.000127)	(0.000116)	(0.000181)
Protest	0.541***	0.257***	0.480***	0.354***	0.485***	0.391***
	(0.139)	(0.0972)	(0.134)	(0.0793)	(0.0958)	(0.0947)
Adj. R-squared	0.381	0.339	0.345	0.330	0.334	0.318
Municipality × month of year FE	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Weather & Protest × base controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,280,752	2,270,197	2,259,648	2,249,105	2,238,560	2,228,016
Municipalities	9,688	9,688	9,688	9,688	9,688	9,688
Mean dep. var.	0.000160	0.000158	0.000158	0.000158	0.000159	0.000159

Note: SE clustered by municipality; *p<0.1; **p<0.05; ***p<0.01. Time horizon is January 2015 until December 2019. Panel A: two-way fixed effects regression with municipality and week fixed effects. Main treatment is measured as the log of 1 + total participants at a PEGIDA Monday demonstration using data from Kanol & Knoesel (2021). Panel B: two-way fixed effects regression with municipality-month of the year and week fixed effects. Main treatment is the interaction between pleasant weather conditions during a Monday and the presence of protest in that week. In this panel, we additionally include weather and protest separately. We also include the interaction of all the controls with the both weather conditions and the presence of protest. For both panels; outcome comes from ProAsyl and AAF and is measured as a dummy variable equals to 1 if any hate crime against refugees was committed between Tuesday and Sunday of the same week (for column 1) or, respectively, 1, 2, 3, 4 and 5 weeks after for columns 2, 3, 4, 5, 6. Municipality-level base controls include: the lagged log (1 +) number of protest participants in the previous week and the lagged dependent variable; GDP per capita; population density; unemployment share; refugee share, latest national vote shares for the AfD and crime rate per 100K inhabitants. Demographic controls include: share of unskilled; share of females; dummies for share of population aged 0-25 25-50 and 50-75. Refugee controls include: share of foreign unemployed; shared of skilled foreigners; share of asylum status granted. Social media controls include: the baseline Twitter penetration at the NUTS-3 level in 2014, the number of tweets containing #refugeesWelcome at the NUTS-3 level per capital in 2014 and the number of Facebook users following AfD before 2015 scaled by population (taken from Müller & Schwarz, 2021) at the groups of municipality level. Columns 5 and 6 present sub-sample analyses with states of the former GDR (Brandenburg, Mecklenburg-Western Pomerania, Saxony, Saxony-Anhalt, and Thuringia) and Western Germany, i.e. Federal Republic of Germany (including Berlin).

Table A.3. **Effect of immigrants characteristics and protest on hate crimes**

	(1)	(2)	(3)	(4)	(5)
Log(1+participants)	0.0442***	0.0430***	0.0436***	0.0547***	0.0459***
	(0.00695)	(0.00715)	(0.00824)	(0.00949)	(0.0106)
× PC1 Economic Threat	0.0175***				
	(0.00489)				
× Refugee share		0.0153***			0.00983***
		(0.00354)			(0.00349)
× Ratio unemployment foreigners/total			0.0163***		0.00287
			(0.00603)		(0.00615)
× Share unskilled among foreigners				0.0216**	0.0100
				(0.00857)	(0.0108)
Interacting variable	-3.39e-05	-7.23e-05	-1.39e-05	1.99e-05	
	(2.26e-05)	(5.82e-05)	(8.45e-06)	(3.74e-05)	
Adj. R-squared	0.340	0.338	0.332	0.337	0.340
Observations	2,280,286	2,280,752	2,280,286	2,280,752	2,280,286
Municipalities	9,686	9,688	9,686	9,688	9,686
Mean dep. var.	0.000160	0.000160	0.000160	0.000160	0.000160
Municipality FE	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes
Base controls	Yes	Yes	Yes	Yes	Yes

Note: SE clustered by municipality; *p<0.1; **p<0.05; ***p<0.01. Time horizon is January 2015 until December 2019. Two-way fixed effects regression with municipality and week fixed effects, and interaction of the main treatment with characteristics at the municipality-week level. All interacted variables are standardized. Column 2 uses the share of refugees in the municipality, column 3 the ratio of the unemployment rate of foreigners over the unemployment rate of locals, and column 4 the share of unskilled workers among foreigners. Column 1 uses as interacting variable the first principal components of the three variables, and column 5 presents a horse-race between the three interacting variables. Main treatment is measured as the log of 1 + total participants at a PEGIDA Monday demonstration using data from Kanol & Knoesel (2021). Outcome comes from ProAsyl and AAF and is measured as a dummy variable equals to 1 if any hate crime against refugees was committed between Tuesday and Sunday of the same week. Municipality-level controls include: the lagged log (1 +) number of protest participants in the previous week and the lagged dependent variable; GDP per capita; population density; unemployment share; refugee share; latest national vote shares for the AfD and crime rate per 100K inhabitants.

Appendix B: Robustness Checks

B.1 TWFE with staggered treatment

Recent developments in the DID literature emphasize caveats in the classical DID setting when it involves both group and time fixed effects, named in the literature Two-Way Fixed-Effects or in short TWFE (e.g., De Chaisemartin & d'Haultfoeuille (2020), Wooldridge (2021), Roth et al. (2022), Goodman-Bacon (2021)). According to this literature TWFE is estimating a weighted sum of several DiDs. When more than two periods and units are treated at different times, the estimator might make "forbidden" comparisons. By "forbidden" we mean comparing units who are both already treated, which might generate negative weights. The negative weights problem is serious since we might get coefficients in the opposite sign of the treatment effects, e.g. negative even if all effects are positive. There are several alternative estimators to tackle the problem, which differ by the specific design that is in use. Some authors also offer a way to diagnose how severe is the problem, for example by finding whether there are negative weights and how many (De Chaisemartin & d'Haultfoeuille (2020)).

Our DiD design is rather complicated in this sense. It is a dynamic non-staggered design (i.e., the treatment is heterogenous across periods and groups, and turns on-and-off), with a non-binary treatment (in the TWFE specification). The reduced form specification is similarly at risk. We first diagnose the TWFE weights using de Chaisemartin et al.'s `twowayfeweights` State command, which reports the sum of positive and negative weights, and the degree of heterogeneity in treatment effects that would be necessary for the estimated treatment effect to have the "wrong sign." If the sum of negative weights is small, and the standard deviation reported is large – then the heterogeneous treatment effect shouldn't be a concern. We present the output of the command for the four sets of controls and the two specifications (TWFE and reduced form) respectively in Panel A and B of Table B.1. In the case of the reduced form specification, we treat the fixed effect structure as a two-way fixed effects structure by considering that observations are identified by a time variable and a pair of municipality and month of year. The TWFE specification, presented in Panel A, has consistently very low negative coefficients. Accordingly, the standard deviation in treatment effect needed to have null results is quite high (around 0.17, three times the effect), and the standard deviation needed to overturn the results is implausibly high (around 14). In such a case, we prefer to use the TWFE estimator since it is more easily interpretable and the most efficient estimator (Wooldridge, 2021).

The Reduced Form specification is more problematic, since the sum of negative coefficients is large. Following Roth et al. (2022) and De Chaisemartin & D'Haultfoeuille (2022b) the appropriate estimator for such designs is De Chaisemartin & D'Haultfoeuille (2022a)'s DID_l estimator. This estimator computes the "weighted average, across time periods t and possible values of the treatment d, of DIDs comparing the t-l-1 to t outcome evolution, in groups with a treatment equal to d at the start of the panel and whose treatment changed for the first time in t-l, the first-time switchers, and in groups with a treatment equal to d from period 1 to t, the not-yet switchers.

DID_l estimates the effect of having switched treatment for the first time l periods ago. The DID_l estimators are unbiased under heterogeneous and dynamic effects", as is our treatment (De Chaisemartin & D'Haultfoeuille, 2022a). We use de Chaisemartin et al.'s `did_multiplegt` Stata package to compute this estimator. The results are presented in Table B.1: while the point estimates are positive and of similar magnitude to the TWFE estimate, the standard errors are high, and only the base controls give a significant result (at the 90% level).

Table B.1. **Estimates robust to negative weights**

	Base controls	+ Demographic controls	+ Refugee controls	+ Social media controls
	(1)	(2)	(3)	(4)
Panel A: Effect of Protest Participation				
TWFE estimate:	0.0631***	0.0631***	0.0630***	0.0630***
Log(participants)	(0.0164)	(0.0164)	(0.0164)	(0.0164)
Sum positive weights	1.0001	1.0001	1.0001	1.0001
Sum negative weights	0.0001	0.0001	0.0001	0.0001
Min $\sigma(\Delta)$ compatible with $\Delta_{TR} = 0$	0.1723	0.1723	0.1721	0.1721
Min $\sigma(\Delta)$ compatible with $\Delta_{TR} < 0$	14.6884	14.7196	14.8174	14.8174
Municipality FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes
Baseline controls	Yes	Yes	Yes	Yes
Demographic controls		Yes	Yes	Yes
Refugee controls			Yes	Yes
Social media controls				Yes
Panel B: Protest on pleasant days				
TWFE estimate:	0.0776***	0.0666***	0.0667***	0.0687***
Pleasant weather × Protest	(0.0219)	(0.0198)	(0.0197)	(0.0203)
Sum positive weights	2.1660	167.4829	60.0944	40.4735
Sum negative weights	-1.1660	-166.4829	-59.0944	-39.4735
Min $\sigma(\Delta)$ compatible with $\Delta_{TR} = 0$	0.0186	0.0001	0.0004	0.0005
Min $\sigma(\Delta)$ compatible with $\Delta_{TR} < 0$	0.0263	0.0001	0.0004	0.0006
DiD estimate:	0.1034*	0.0933	0.0763	0.0812
Pleasant weather × Protest	(0.0583)	(0.0879)	(0.236)	(0.473)
Municipality × month of year FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes
Weather & Protest × baseline controls	Yes	Yes	Yes	Yes
Weather & Protest × demographic controls		Yes	Yes	Yes
Weather & Protest × refugee controls			Yes	Yes
Weather & Protest × social media controls				Yes
Observations	2,156,734	2,156,734	2,156,734	2,156,734
Municipalities	9,023	9,023	9,023	9,023
Mean dep. var.	0.000168	0.000168	0.000168	0.000168

Note: Panel A shows the TWFE coefficients, and the sum of negative weights computed using the `twowayfeweight` command. Panel B shows the same for the Reduced Form estimate, and additionally shows the DiD_l estimate of De Chaisemartin & D'Haultfoeuille (2022a).

B.2 Other robustness checks

We conduct other robustness checks on our main specification presented in Table B.2, Table B.3 and Table B.4. First, we vary the fixed-effects specifications. presents various alternative specifications. Additionally, we conduct robustness checks, presented in Table B.3, that restrict the sample; aggregate to a higher regional level, change the estimation method and add additional set of controls. Finally we allow our observation to be spatially correlated at different distance windows. For the three tables, column 1 presents the baseline specification for comparison.

Variation on the fixed effect structure Our main analysis include municipality and week fixed effects. We show in Table B.2 that the results are robust to the inclusion of more demanding fixed effects. Panel A corresponds to the TWFE identification and panel B to the reduced form. Column 1 uses only municipality and week fixed-effects (corresponding to our main TWFE specification), column 2 replaces the municipality fixed effects by fixed effects of the municipality and month of year pair (this corresponds to our main reduced form specification). Column 3 adds per-municipality time trends to the basic specification. Column 4 includes municipality fixed effects and fixed effects per NUTS-1 and week combination. Column 5 combines these with municipality month of year fixed effects, and column 6 with per-municipality trends. The results do not lose significativity under these richer fixed effects structures. The addition of NUTS-1 week effects slightly lowers the magnitude of the TWFE estimates.

Changing the geographical aggregation We chose to do our analysis at the municipality level, which is the finest geographical level at which data is available to us. However, the delimitation of municipalities and districts changed during our study period and many of our controls are only available at higher geographical levels. Moreover, there could be spillover and displacement effects to and from surrounding municipalities influencing our results. In columns 2 and 3 of Table B.3, we collapse our dataset to higher geographical administrative levels (NUTS-2 and NUTS-3 instead of districts for column 2 and 3 respectively). The collapsed results are close to the results at municipality level, with the NUTS-3 results being almost equal (for the TWFE specification) or slightly lower (for the reduced form model), and the NUTS-2 results being slightly higher than the baseline results.

Restricting the sample to municipalities with protest In our main specification we consider all municipalities (the more granular geographic administrative level). However, a lot of municipalities are very small and have a very low population. In those municipalities both protest and hate crimes are very unlikely. We then restrict the sample to municipalities that had at least a protest during the period of analysis to avoid an over-representation of zeros (column 4 of Table B.3). Results remain positive and significant with a magnitude slightly higher,

Logit and probit Our main specification uses a linear probability model. The effect of municipality and day fixed effects might be multiplicative instead of linear: for instance, a given event in Germany may double the baseline probability of hate crimes in all municipalities instead of adding a constant. We check that our results are robust to using non-linear models. First, we limit our dataset to municipalities where at least one hate crime occurred, and days where at least one hate crime occurred (in other words, to observations whose result is not fully determined by the fixed effects). Column 5 Table B.3 presents the result of estimating the linear probability model on this subsample: the effects are higher for both specifications, and still significant. Next, we estimate both a logit model (column 6) and probit model (column 7). The results stay positive and significant, both for the TWFE and reduced form empirical strategy (the magnitudes are not directly comparable between models).

Controlling for weather during hate crimes Crime rates are known to be influenced by weather (Chersich et al., 2019), and weather is temporally correlated. Thus, in our reduced-form specification, part of the crime rates during the week might be explained by weather on Monday. To check that this does not influence our results, we change our specification to consider hate crimes at the daily level, with one observation per municipality and day of the week (Tuesday to Sunday) instead of one observation per week, and control by whether the weather was pleasant on the day where a hate crime could have been committed. Column 8 Table B.3 presents baseline results for the daily specification for comparison, and column 9 adds the weather control. The coefficient of weather is very small, and no effect on the results is visible.

Controlling for past cumulative protest and hate crimes Our main specification controls for the number of protest and hate crimes the week before. However, events happening even before may also potentially affect the protest and the likelihood of hate crimes and can lead to an overestimation of the effect. We show in column of Table B.3 that our results hold when including the total cumulative number of past protest and the total cumulative number of past hate crimes measured in the week before the outcome.

Allowing for spatial correlation Observations can be correlated spatially which may lead to a wrong estimation of the standard error. We allow the spatial correlation of our observation within certain spatial windows using Conley standard errors (Conley, 1999). Table B.4 presents the results. Panel A corresponds to the TWFE identification and panel B to the reduced form. For the TWFE estimation, our results remain strongly significant (at 1% level) when allowing correlation within 25km, 50km, 100km or 200km. For the reduce form estimation, the results lose significance and slightly pass the threshold of 10% (p_values from 0.126 and 0.134). For computational reasons, sample is restricted to municipalities that had at least one protest during our period of study.

Table B.2. **Robustness: varying the fixed effects structure**

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Protest Participation						
Log(1+participants)	0.0524***	0.0538***	0.0512***	0.0424***	0.0427***	0.0412***
	(0.0102)	(0.0108)	(0.0101)	(0.00815)	(0.00824)	(0.00782)
Adj. R-squared	0.325	0.329	0.334	0.126	0.137	0.134
Base controls	Yes	Yes	Yes	Yes	Yes	Yes
Panel B: Protest on pleasant days						
Pleasant weather × Protest	0.0773***	0.0798***	0.0768***	0.0746**	0.0753**	0.0747**
	(0.0265)	(0.0260)	(0.0259)	(0.0311)	(0.0309)	(0.0304)
Weather	0.000145	2.60e-05	0.000132	0.000194*	4.27e-05	0.000202*
	(9.94e-05)	(9.28e-05)	(0.000103)	(0.000105)	(0.000105)	(0.000107)
Protest	0.506***	0.526***	0.472***	0.469***	0.480***	0.440***
	(0.137)	(0.136)	(0.135)	(0.154)	(0.148)	(0.152)
Adj. R-squared	0.372	0.379	0.376	0.154	0.165	0.157
Weather × base controls	Yes	Yes	Yes	Yes		
Protest × base controls	Yes	Yes	Yes	Yes		
Observations	2,280,752	2,280,752	2,280,752	2,280,281	2,280,281	2,280,281
Municipalities	9,688	9,688	9,688	9,686	9,686	9,686
Mean dep. var.	0.000159	0.000159	0.000159	0.000159	0.000159	0.000159
Municipality FE	Yes		Yes	Yes		Yes
Week FE	Yes	Yes	Yes			
Municipality × month of year FE		Yes			Yes	
Per-municipality time trend			Yes			Yes
Week - NUTS-1 FE				Yes	Yes	Yes

Note: SE clustered by municipality; *p<0.1; **p<0.05; ***p<0.01. Time horizon is January 2015 until December 2019. Panel A: two-way fixed effects regression with municipality and week fixed effects. Main treatment is measured as the log of 1 + total participants at a PEGIDA Monday demonstration using data from Kanol & Knoesel (2021). Panel B: two-way fixed effects regression with municipality-month of the year and week fixed effects. Main treatment is the interaction between pleasant weather conditions during a Monday and the presence of protest in that week. In this panel, we additionally include weather and protest separately. We also include the interaction of all the controls with the both weather conditions and the presence of protest. For both panels; outcome comes from ProAsyl and AAF and is measured as a dummy variable equals to 1 if any hate crime against refugees was committed between Tuesday and Sunday of the same week. Municipality-level controls include: the lagged log (1 +) number of protest participants in the previous week and the lagged dependent variable; GDP per capita; population density; unemployment share; refugee share; latest national vote shares for the AfD and crime rate per 100K inhabitants.

Table B.3. **Robustness checks**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	baseline	collapsed		only	only with hate crimes			daily data		cumul.
		NUTS-2	NUTS-3	treated	linear	logit	probit	baseline	weather control	control
Panel A: Protest Participation										
Log(1 + participants)	0.0524***	0.0549***	0.0530***	0.0574***	0.0678***	0.870***	0.440***	0.0152***	0.0152***	0.0525***
	(0.0102)	(0.0127)	(0.0102)	(0.0104)	(0.00931)	(0.0578)	(0.0266)	(0.00512)	(0.00512)	(0.0101)
Weather (crime)									-3.51e-06	
									(4.65e-06)	
Adj. R-squared	0.325	0.348	0.327	0.429	0.378			0.172	0.172	0.330
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes		Yes
Day FE									Yes	
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Panel B: Protest on pleasant days										
Weather × Protest	0.0798***	0.0888**	0.0733***	0.0814**	0.0989*	1.234*	0.788**	0.0170*	0.0170*	0.0805***
	(0.0260)	(0.0337)	(0.0249)	(0.0301)	(0.0540)	(0.696)	(0.326)	(0.0102)	(0.0102)	(0.0266)
Weather (protest)	2.60e-05	-0.0232	0.000657	0.0515*	-0.0728	-0.994	-0.499	1.44e-05	1.45e-05	5.99e-05
	(9.28e-05)	(0.0170)	(0.00155)	(0.0268)	(0.0893)	(1.388)	(0.672)	(2.92e-05)	(2.91e-05)	(0.000110)
Protest	0.526***	-0.225	0.453***	0.840***	1.183***	14.56**	8.224**	0.193***	0.193***	0.539***
	(0.136)	(0.213)	(0.138)	(0.172)	(0.396)	(6.511)	(3.311)	(0.0439)	(0.0439)	(0.138)
Weather (crime)									-4.98e-07	
									(4.48e-06)	
Adj. R-squared	0.379	0.404	0.383	0.491	0.353			0.213	0.213	
Municipality × month of year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes		Yes
Day FE									Yes	
Weather & Protest × baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,280,752	9,571	99,941	6,447	14,804	14,804	14,804	13,684,512	13,684,512	2,280,752
Municipalities	9,688			25	84	84	84	9,688	9,688	9,688
NUTS		37	286							
Mean dep. var.	0.000160	0.0370	0.00363	0.0425	0.0246	0.0246	0.0246	4.65e-05	4.65e-05	0.000160

Note: SE clustered by municipality (NUTS-2 in col. 2, NUTS-3 in col. 3); *p<0.1; **p<0.05; ***p<0.01. Time horizon is January 2015 until December 2019. Panel A: two-way fixed effects regression with municipality and week fixed effects. Main treatment is measured as the log of 1 + total participants at a PEGIDA Monday demonstration using data from Kanol & Knoesel (2021). Panel B: two-way fixed effects regression with municipality-month of the year and week fixed effects. Main treatment is the interaction between pleasant weather conditions during a Monday and the presence of protest in that week. In this panel, we additionally include weather and protest separately. We also include the interaction of all the controls with the both weather conditions and the presence of protest. For both panels; outcome comes from ProAsyl and AAF and is measured as a dummy variable equals to 1 if any hate crime against refugees was committed between Tuesday and Sunday of the same week. Municipality-level base controls include: the lagged log (1 +) number of protest participants in the previous week and the lagged dependent variable, GDP per capita; population density; unemployment share; refugee share, latest national vote shares for the AfD and crime rate per 100K inhabitants. Col. 1 presents the baseline specification. In Col. 2, observations at at the NUTS-2 level (NUTS-3 in Col. 3). Col 4 considers only treated municipalities (ie. with at least one protest). Col 5-7 restrict to FE units (municipalities and day, or municipality-months and days) with at least one hate crime. Col 5 uses the usual linear specification, col 6 and 7 use respecitvely logit and probit specifications. Col 8 and 9 use data on hate crimes at the daily level instead of weekly: there is one observation per day of the week (excluding Monday). Col 9 additionally controls for whether the weather was pleasant on the day the hate crime would have occurred. Col. 10 controls by the cumulative number of protests and hate crimes until now.

Table B.4. **Conley standard errors**

	(1)	(2)	(3)	(4)
Panel A: Protest Participation				
Log(1 + participants)	0.0574***	0.0574***	0.0574***	0.0574***
	(0.00530)	(0.00534)	(0.00557)	(0.00601)
Adj. R-squared	0.309	0.309	0.309	0.310
Municipality FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes
Baseline controls	Yes	Yes	Yes	Yes
Panel B: Protest on pleasant days				
Weather × Protest	0.0814	0.0814	0.0814	0.0814
	(0.0532)	(0.0533)	(0.0538)	(0.0543)
Weather	0.0515	0.0515	0.0515	0.0515*
	(0.0318)	(0.0313)	(0.0320)	(0.0310)
Protest	0.840***	0.840***	0.840***	0.840***
	(0.167)	(0.168)	(0.171)	(0.175)
Adj. R-squared	0.377	0.380	0.381	0.383
Municipality × month of year FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes
Weather & Protest × baseline controls	Yes	Yes	Yes	Yes
Observations	6.447	6.447	6.447	6.447
Municipalities	25	25	25	25
Mean dep. var.	0.0425	0.0425	0.0425	0.0425
Cutoff distance	25 km	50 km	100 km	200 km

Note: Conley spatially clustered SE, with cutoff distance 25, 50, 100 and 200 km for columns 1, 2, 3, 4; *p<0.1; **p<0.05; ***p<0.01. Sample is restricted to treated municipalities, ie. with at least one protest. Time horizon is January 2015 until December 2019. Panel A: two-way fixed effects regression with municipality and week fixed effects. Main treatment is measured as the log of 1 + total participants at a PEGIDA Monday demonstration using data from Kanol & Knoesel (2021). Panel B: two-way fixed effects regression with municipality-month of the year and week fixed effects. Main treatment is the interaction between pleasant weather conditions during a Monday and the presence of protest in that week. In this panel, we additionally include weather and protest separately. We also include the interaction of all the controls with the both weather conditions and the presence of protest. For both panels; outcome comes from ProAsyl and AAF and is measured as a dummy variable equals to 1 if any hate crime against refugees was committed between Tuesday and Sunday of the same week. Municipality-level base controls include: the lagged log (1 +) number of protest participants in the previous week and the lagged dependent variable, GDP per capita; population density; unemployment share; refugee share, latest national vote shares for the AfD and crime rate per 100K inhabitants.

Appendix C: Data Appendix

C.1 Details on Protest Data

To create this dataset, the Kanol & Knoesel (2021) identified relevant parliamentary questions that contained information on right-wing extremist demonstrations. They then extracted the relevant data from tables included in these responses and merged them to create a comprehensive dataset. To classify each demonstration based on its ideology, an identification variable was added to the dataset. This classification process was based on descriptions provided in the government's responses to parliamentary questions. Demonstrations were classified as "right-wing extremist," "mostly right-wing," or "partially right-wing" based on these descriptions. The number of right-wing protests was highest in 2015 (with 290 demonstrations) and lowest in 2010 (with only 70 demonstrations). Of all demonstrations in this dataset, over 83% were classified as "right-wing extremist," while around 17% were categorized as "mostly right-wing." Only a very small fraction (0.2%) was identified as "partially right-wing."

The authors also used geocoding techniques to identify the location of each demonstration. This involved converting textual descriptions of locations into geographic coordinates that could be plotted on a map. Some demonstrations were held in more than one place or moved through multiple locations. We treat these protests as separate incidents. In some cases, exact numbers for participants in a demonstration were not available; instead, an estimation was given (e.g., 5-10 or 100-500). In these cases, we follow the authors and use the average of this range of numbers.

C.2 Data coherence

Our regional-level of analysis is the municipality level as defined above. However some variables are available only at the district level, the values of these variables remain identical across municipalities within the same district. Both municipalities' and districts' borders changed during our sample period, i.e., 2015 - 2020. Hence, for reasons of data coherence, we adjusted all variables according to one border division.

Municipality-level variables are adjusted to the border division as of December 2020. We use the 'name and area changes of municipalities' tables which are published yearly by DESTATIS, which document four types of municipality changes: (1) municipalities that merged with other municipalities, or joined an existing municipality; (2) municipalities that split to several municipalities; (3) change of key; (4) change of name. Municipality-level variables from before December 2020 were updated as follows. First, names and keys were updated to December 2020. Second, for merged municipalities (i.e., change (1)), averaged variables (e.g., voting turnout) were updated as a population-weighted average, for municipality i in year t=2015,...,2020, and the new merged municipality j: $Var_{j,t} = \sum_{i=1,\ldots,n}^{n} Var_{i,t} * \frac{Pop_{i,t-1}}{Pop_{j,t}}$; and for summed variables (e.g., total votes for AfD) $Var_{j,t} = \sum_{i=1,\ldots,n}^{n} Var_{i,t}$. The split municipalities (i.e., change (2)) were dropped due to their small share in the sample, with less than half a percent of all municipalities.

District-level variables are adjusted to the NUTS3 2013 version, which entered into force on

31 December 2013 and applied from 1 January 2015. Two changes were made since then in 2017: (1) The border between Cochem-Zell and Rhein-Hunsrück-Kreis slightly shifted, without affecting other districts borders; (2) Göttingen and Osterode am Harz merged into one district under the name of Göttingen. Change (1) was ignored, since the boundary shift is minor in terms of km^2 area. To account for change (2), all regional controls in Göttingen and Osterode am Harz after 2016, received a value equal to the value in Göttingen, weighted by the share of the corresponding district in 2016, s.t.: for averaged variables (e.g., unemployment rate) $Var_{i,t} = Var_{i,t} * \frac{Var_{i,2016}}{Var_{j,2016}}$; and for summed variables (e.g., population) $Var_{i,t} = Var_{i,t} * \frac{Var_{i,2016}}{Var_{i,2016}+Var_{j,2016}}$, with $i, j \in$ [Göttingen, Osterode am Harz] and $t \in [2017, 2020]$ Moreover, GDP data was available only in the NUTS3 2016 format (i.e., also before 2017). Hence only GDP was recovered weighted by the population share of the two regions.

Table C.1. Description of Data Sources

Variable	Regional Level	Period	Source
Main variables			
Participants in PEGIDA protests	muni	2015-2020	Kanol & Knoesel (2021)
Hate crimes	muni	2015-2020	Amadeu Anotonio Foundation and PRO ASYL Foundation
Weather	muni	2015-2019	European Centre for Medium-Range Weather Forecasts (ECMWF)
Base controls			
GDP per capita	dist	2000-2019	Federal Statistics Office
Population density	muni	2009-2021*	own calculations
Unemployment rate	muni	2008-2021**	Federal Employment Agency
AfD votes (Bundestag)	muni	2013-2021***	Federal Returning Officer
Total crime cases (per 100k pop)	muni	2013-2021	Federal Criminal Police Office
Demographic controls			
Workers without qualification	dist	2008-2021****	Federal Statistics Office
Age distribution (0-25, 25-50, 50-75)	muni	2009-2021*	Federal Statistics Office
Gender distribution	muni	2009-2021*	Federal Statistics Office
Immigration-related controls			
Unemployment rate of non-Germans	muni	2008-2021**	Federal Employment Agency
Foreign workers with academic qualification	dist	2008-2021****	Federal Statistics Office
Asylum recipients (share)	dist	2011-2021*	Federal Statistics Office
Social media controls			
Twitter usage per capita	dist	baseline	Twitter API
#RefugeesWelcome tweets per capita	dist	baseline	Twitter API
AfD followers	muni group	baseline	Müller & Schwarz (2023)
Additional variables			
Tweets mentioning PEGIDA	dist	2015-2020	Twitter API

Note: This table provides information of the variables we use for the analysis. The first column describes the geographical level at which we observe each variables. Districts (402 of them) are equivalent to NUTS3, Municipalities are smaller. Column 2 describes the period for which we have information of each variable. $*$ stands for up until 31.12 of the previous year; $**$ indicates that we have the information on the yearly/monthly average. $***$ during this period, every election $****$ as of 30.06 of the year. Column 3 provides the source from which we extract each of the variables.