2019

The Effects of Dog-Whistle Politics on Political Violence

Olga Chyzh
Iowa State University, ochyzh@iastate.edu

Mark David Nieman
Iowa State University, mdnieman@iastate.edu

Clayton Webb
University of Kansas

The Effects of Dog-Whistle Politics on Political Violence

Abstract
The election of President Trump marked significant changes in the content, outlets, and the level of civility of political rhetoric. The traditional left/right policy disagreements took on a more populist tone, activating extremist elements within society. We explore the consequences of political appeals to nationalist identity within the context of modern-day America. We argue that employed by elected officials, nationalist political rhetoric legitimizes extremist views and their expression. This effect is exacerbated by the social media, which provides an unmoderated channel for communication between elected officials and their extremist supporters. We test the link between nationalist rhetoric and hate crimes using data collected from Twitter, as well as an original dataset on daily hate incidents in the US, between February 2017–April 2018, and find strong evidence for our theory. Our results have important implications for the study of political communication and political violence.

Disciplines
Political Science

Comments
This Unpublished article is cited as Chyzh, O.V., Nieman, M.D., Webb, C., The Effects of Dog-Whistle Politics on Political Violence. 2019; 1-10. Posted with permission.

The Effects of Dog-Whistle Politics on Political Violence

Olga V Chyzh[*] Mark Nieman[†] Clayton Webb[‡]

May 20, 2019

Abstract

The election of President Trump marked significant changes in the content, outlets, and the level of civility of political rhetoric. The traditional left/right policy disagreements took on a more populist tone, activating extremist elements within society. We explore the consequences of political appeals to nationalist identity within the context of modern-day America. We argue that employed by elected officials, nationalist political rhetoric legitimizes extremist views and their expression. This effect is exacerbated by the social media, which provides an unmoderated channel for communication between elected officials and their extremist supporters. We test the link between nationalist rhetoric and hate crimes using data collected from Twitter, as well as an original dataset on daily hate incidents in the US, between February 2017–April 2018, and find strong evidence for our theory. Our results have important implications for the study of political communication and political violence.

[*] Departments of Political Science and Statistics, Iowa State University
[†] Department of Political Science, Iowa State University
[‡] Department of Political Science, University of Kansas

Introduction

The election of President Trump marked significant changes in the content, outlets, and the level of civility of political rhetoric in the US. The traditional left/right policy disagreements took on a more populist tone (Boucher and Thies, 2019; Jamieson and Taussig, 2017). The President's norm-shattering rhetoric, combined with his preference for communicating via Twitter, rather than more conventional news outlets, created space for expression of more radical and marginalized views with notable white nationalist undertones (Barkun, 2017). As leader of the Republican party, Trump's political positions and rhetoric have been adopted by many down-ballot co-partisans (Bartels, 2018; Jacobson, 2018). The distinction between present-day Republican rhetoric and overt hate speech has become so blurred that Twitter is unable to identify and ban white nationalist content using the same algorithms used to ban ISIS accounts out of fear of accidentally removing some tweets by the President and other politicians (Panetta, 2019).

We argue that the normalization of previously marginalized views and rhetoric, especially rhetoric that casts society in terms of in-groups and out-groups, also normalizes other types of hate expression against the perceived out-groups, e.g., threats or even violent actions. Proponents of radical views take cues from political elites, especially elected government officials. Social media outlets, such as Twitter and Facebook, reinforce this effect by providing channels for direct and unmoderated communication and immediate feedback, and by facilitating communication among like-minded individuals. While radical political views may remain marginalized within the broader population, observing supportive statements from elected officials, as well as the ability to connect to other like-minded individuals on social media, may exaggerate the perceived societal support for extremist views.

Legitimization of extremist nationalist views activates previously latent extremist audiences into following and reacting to political rhetoric, both by the elected officials they perceive as sympathizers and by those who condemn their views. Once activated, extremists interpret all of the actions of their perceived sympathizers as additional validation for their views and expression. Simultaneously, rhetoric by politicians who condemn extremism may create a 'lashing-out' effect.

We explore the relationship between political rhetoric, especially that aimed at extremist audiences, and political violence, using the data on the Twitter activity by the members of the United States (US) House of Representatives and an originally collected dataset on hate crime incidents in the US between February 15, 2017 and April 5, 2018. We identify the US representatives who are most likely to engage audiences with anti-minority views, based on the proportion of self-identified white nationalist followers of each representative. We find a strong correlation between tweets by this subset of US representatives and incidents of hate crimes, ranging from racially charged instances of vandalism and threats to physical assault and murder. Our analysis also shows that tweets by these representatives are associated with incidents of hate crimes, irrespective of the representatives' political party.

This study is among the first to explore the consequences of social media political engagement beyond the sphere of social media. Our results suggest that seemingly harmless political engagement with extremist groups is, in actuality, neither harmless nor inconsequential. More broadly, our results highlight the possible risks associated with direct and unmoderated social media engagement between politicians and their supporters.

The rest of the paper is organized as follows. In the next section, we review the literature and present the theoretical argument that links direct political engagement and hate crimes. Next, we describe the research design and the procedures used to collect both Twitter data as well as original data on hate crimes in the US. Finally, we present and discuss the results of the statistical tests, including a series of robustness checks, and conclude by summarizing the contribution and highlighting pathways for future research.

Dog-Whistle Politics and Race

Political elites have long relied on the media outlets, such as television, radio, or newspapers, to share or amplify their messages to their constituents. First the candidacy, and then the election, of President Trump marked a shift from political engagement via these traditional media outlets to a qualitatively different type of political engagement via social media, such as Twitter or Facebook. In contrast to traditional media, where the political message is often moderated as it goes through the filter of editing and production, messages shared via social media are instantaneous, direct, and often unedited. By providing a platform for these types of messages, social media outlets enable explicitly inflammatory or politically divisive rhetoric, as well as multivocal and coded appeals to more extremist groups of supporters.

Multivocal and dog-whistle appeals have long been important forms of political communication. Multivocal appeals are single statements that are intended to be interpreted differently by different audiences, where the out-group is generally oblivious to the in-group meanings (Padgett and Ansell, 1993; Tilly, 2003). That the out-group does not infer the same meaning as the in-group is the key to the message's success, as an explicit appeal may incur a costly negative reaction by the out-group. An experimental study by Albertson (2015) uses national samples from religious (in-group) and non-religious (out-group) populations to find that, while both multivocal and explicit religious appeals were persuasive to in-group respondents, the former were unnoticed by out-group respondents and the latter evoked negative reactions among out-group respondents.

In contrast to multivocal appeals, coded or dog-whistle appeals do not require that the out-group fail to understand the appeal, only that the speaker have some level of plausible deniability of the messages true meaning (Mendelberg, 2001; Valentino, Hutchings and White, 2002; White, 2007). Examples of a dog-whistle appeal include the 1988 Willie Horton ad ran against Democratic presidential candidate, Michael Dukakis, and references to the "inner city" when discussing anti-crime policies (Hurwitz and Peffley, 2005). Dog-whistle appeals are frequently invoked by political elites when pursuing economic policies that advantage economic elites over the middle class, as a means to split coalitions and make race and ethnicity more salient than bread-and-butter issues (López, 2015).

Multivocal appeals rely on a positive reaction among the in-group and minimal reaction among the out-group; dog-whistle appeals intend to create an 'us' versus 'them' dynamic where the gains among the former outweigh the losses for the latter. The rationale for success in using dog-whistle appeals relies on the basic psychology of individuals and their propensity to identify themselves with, and sort themselves into, social groups (Tajfel, 1981; Brewer, 1999, 2007; Fearon and Laitin, 2000; Kinder and Kam, 2010). While individuals have a number of social identities that they can draw on, political elites often attempt to emphasize some of these identities more than the others (Quillian, 1995; Gurr, 2000; Penn, 2008; Wood, 2008).

The symbolic politics literature shows that individuals are more likely to adopt specific identities (e.g., with their party, liberal or conservative ideology, nationalism, or racial prejudice) in response to affective triggers, and that political elites employ such triggers to maximize support for their platforms (Sears, Hensler and Speer, 1979; Sears et al., 1980). One branch of this research—the study of symbolic racism—links racial and ethnic dog-whistles to individual's attitudes and behavior (Sears and Kinder, 1971; Kinder and Sanders, 1996; Feldman and Huddy, 2005). The central idea is that the racial or ethnic majority rejects the belief that the minority continues to suffer from an acknowledged history of discrimination, and becomes convinced that the minority group violates accepted national values (Kinder and Sears, 1981). A number of studies have shown that symbolic racism is a key predictor of political attitudes towards crime, affirmative action, and welfare (for a review of these studies, see Sears and Henry, 2005). Recent work has shown that explicit symbolic racism continues to be a better predictor of public opinion variables than implicit measures (Ditonto, Lau and Sears 2013; see also Hughey and Parks 2014).

Dog-whistle appeals to symbolic racism continue to be effective, in part due to a rise in so-called resentment politics or the majority belief that government programs and policies intended to prevent discrimination of racial and ethnic minorities are actually the cause of reverse discrimination against the majority. Norton and Sommers (2011) show evidence of a rise in resentment politics, finding that both white and black respondents believe that anti-black discrimination have decreased in the US over the last several decades, but white respondents also believe that anti-white bias has increased in proportion with this decline whereas black respondents do not believe this. While political elites have long used resentment politics (López, 2015), Hughey and Parks (2014) argue that its use has seen a uptick following the inauguration of President Obama.

The finding that whites increasingly believe they are victims of reverse discrimination matters because of the known relationship between psychological stress and support for exclusionary attitudes. While there is strong evidence that psychological stress induced from direct exposure to physical threats leads to exclusionary attitudes (Canetti-Nisim et al., 2009), there is also experimental evidence that that this effect is stronger in response to a threat against the members of one's perceived in-group, often defined by ethnic and racial markers (Avdan and Webb, 2019). Support for exclusionary actions in response to perceived threats, however, may be mediated by political ideology (Jungkunz, Helbling and Schwemmer, 2019).

Some recent evidence also shows an increase in effectiveness of explicit racial appeals, with overly hostile messages towards racial and ethnic minorities. A survey experiment study of four nationally representative samples finds that both coded and explicit appeals have equally large and stable effects on racial attitudes (Valentino, Neuner and Vandenbroek, 2018). These findings suggest that, whereas in the past, implicit and symbolic racial appeals were rewarded while explicit appeals were condemned, overt racially hostile messaging no longer elicits either surprise or punishment from the public.

Perceptions of threat and support for exclusionary attitudes are, of course, not unique to the study of US politics. The notion of populist and nationalist bidding is central to theories of diversionary war (Mitchell and Prins, 2004; Pickering and Kisangani, 2005), arguments that states undergoing democratic transitions are more war-prone than other states (Ward and Gleditsch, 1998; Mansfield and Snyder, 2002), and domestic targeting of minorities (Mousseau, 2001; Tir and Jasinski, 2008). Studies using survey data find that respondents from a variety of countries are more supportive of nationalist and exclusionary policies when the question is framed in terms of a threat from an out-group (Gibler, Hutchison and Miller, 2012; Miller, 2017).

Rhetoric and Political Violence

There is growing evidence that violent rhetoric is linked to violent behavior. For example, in a study of hate groups in the state of New York, Asal and Vitek (2018) find that violent rhetoric is the best predictor of group-level violence, while factors such as ideology have little or no effect. This is consistent with experimental work by Kalmoe (2014), who found a positive relationship between violent metaphorical language and support for political violence.

Hateful rhetoric against political minorities seems to be especially dangerous, as concerns about "cultural threats" are known to be associated with violence against out-groups. For example, a study relying on semi-structured interviews with participants of white nationalist internet chat rooms revealed that issues like interracial marriage and integration of neighborhoods are much more likely to lead to calls for violence than issues of job competition with racial minorities (Glaser, Dixit and Green, 2002). Likewise, when comparing a sample from the general population to that of white supremacists and hate crime perpetrators from North Carolina, Green, Abelson and Garnett (1999) found that, despite similar economic conditions, the latter two groups were more opposed to interracial marriage, integration of neighborhoods, and banning the Confederate flag.

Social media has also created more space for hate groups and other proponents of exclusionary policies to share messages and spread propaganda, as well as facilitate militia and lone wolf training, such as sharing bomb making instructions. Chan, Ghose and Seamans

(2015) find that increased access to broadband internet is associated with an increase in the frequency of hate crimes. Further, these effects are stronger in areas with higher pre-existing levels of racism. Rather than the increase in hate crimes resulting from more hate group activity, they appear to result from an increase in the number of lone wolf attacks (Chan, Ghose and Seamans, 2015). From a policy perspective, lone wolf attacks are especially problematic because they are difficult for law enforcement to prevent and are more lethal than other terrorist attacks (Phillips, 2017).[1]

Media Amplification

Media amplification of hostile racial and ethnic rhetoric further increases the likelihood of political violence, especially against minority out-groups. First, mass media outlets have traditionally played a key role in setting the agenda, particularly regarding political narratives (Iyengar and Kinder, 2010), and attitude formation (Zaller, 1992). Through these two mechanisms, media help set initial predisposition and demarcate what are or are not acceptable public positions and behaviors. As we later argue, the public acceptability of positions and attitudes is directly related to the propensity for political violence.

Although controversial statements by government officials and high-profile political candidates certainly garnered media attention, in the past both the media and political elites in liberal democracies tended to condemn or mute any overt racially charged political rhetoric. More recently, radical fringe groups have been able to affect traditional mass media outlets, especially following major political events, such the 9/11 terrorist attacks. Studies have shown that mainstream media reports that involve fearful or angry discourse espoused by fringe groups may give disproportionately large representation to these minority views within the broader political discourse. Comparing 1084 press releases about Muslims by 120 civil society organizations—only a small proportion of which held anti-Muslim biases—to 50407

[1]It is worth noting that acts of actual violence may cause some individuals to no longer support hate organizations. For example, Barceló and Labzina (2018) show that the Islamic State's support on social media, such as Twitter, substantially decreased following major terrorist attacks.

newspaper articles and TV transcripts from 2001 to 2008, Bail (2012) shows that views of a small number of fringe groups have become overrepresented in the mainstream media. One explanation is that the reporters seek to present "both sides" of an issue to appear unbiased.[2]

Another source of media amplification may arise from an increase in the number and types of media outlets. The proliferation of both traditional media outlets, as well as the rise of social media, have effectively ended the monopoly on agenda setting that was once held by a few traditional media outlets. Individuals can now tune into media outlets perceived to hold similar partisan leaning, leading to more polarized news audiences (Iyengar and Hahn, 2009). Moreover, social media enables individuals to largely bypass media agenda-setting efforts altogether. In fact, there is some evidence that the reverse is true—that social media affects the agenda of what stories mass media outlets cover (King, Schneer and White, 2017).

Information and communication technologies are thought to affect collective violence by influencing the marketplace of ideas (Warren, 2015). Whereas traditional mass media cuts across societal cleavages and provides a common base of news and facts, social media increases societal divisions by encouraging horizontal transfers of information. To investigate this mechanism, Warren (2015) looks at a sample of 24 African states, and finds that patterns of political violence increase in areas where social media is more prevalent than mass media. These results are consistent with those of Garcia and Wimpy (2016), who find that increases in cell-phone and internet usage amplify the spread of anti-government violence in 44 African states from 2000–2011.

Government and elected officials use social media to rally support and advance political ends. During the 2014 anti-Maduro protests in Venezuelan, pro-Maduro legislators used Twitter to pose competing narratives in order to obfuscate criticisms by opposition officials (Munger et al., 2018). Likewise, Chinese government officials frequently counter online criticism by shifting the focus to other issues (King, Pan and Roberts, 2017). More recently, as part of an effort to unify their political support, the Myanmar military created fake Face-

[2] As part of the effort to report both sides, reporters may also be "trolled"—intentionally mislead—by actors seeking to spread hateful messages (Phillips, 2018).

book accounts to spread false rumors against the Muslim Rohingya minority (Mozur, 2018). Likewise, the government of Sri Lanka used Facebook to promote fears of a Muslim plot against the Sinhalese majority (Taub and Fisher, 2018).

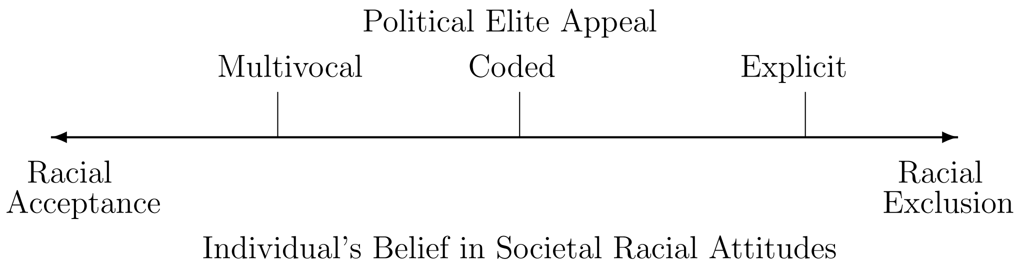
Linking Political Elite Rhetoric and Hate Crimes

Political elites use dog-whistles and heated rhetoric, often with seemingly innocuous intent, e.g., to excite their political base. The close link between political rhetoric and violence, however, is one of the central themes of research on symbolic politics and political rhetoric. Bridging this research with the growing literature on social media, we argue that bringing anti-minority rhetoric into the political mainstream increases the number of incidents of political violence against minority groups. This effect, moreover, is exacerbated by the growth of social media.

The key premise behind our theoretical argument is that the democratic election process (e.g., majority vote) gives elected officials the power and responsibility to shape political norms. By the virtue of holding a political office, government officials also hold the power to define the line between acceptable and unacceptable rhetoric and expression. A public official's use of dispassionate language, and appeals to political moderation and compromise, help create the norms of political tolerance and inclusion, at the same time marginalizing divisive and extremist views. In contrast, an elected official's use of inflammatory, divisive or anti-minority rhetoric, undermines these norms, while validating and normalizing radical political views.

Government officials' ability to shape or undermine norms is both magnified and accelerated by the use of social media. By removing the filter associated with more conventional media engagement, social media provides politicians with two-way, direct, unmoderated channels of communication with their constituents. Importantly, while extremist groups (by definition) make up only a small number of each elected official's constituents, such groups

Figure 1: Political Appeals and Activating Latent Racial Extremists



are substantially more likely to take advantage of these channels of communication (due to the intensity of their beliefs). Once completely isolated within the broader society, individuals with extreme beliefs can now use social media to identify with and form larger communities, which by itself may exaggerate their perception of support for their beliefs within the broader society. Even weak validation by an elected official amplifies this effect, creating a false perception that the extreme political beliefs are actually not that far from the political mainstream.

Moreover, once a politician is associated with radicalized and previously marginalized views, all of their political actions are tinged with this viewpoint. Seemingly mundane political actions, such as introducing legislation in committee or naming government buildings, are actions being done by someone associated with radicalized views. By continuing to serve and participate in the body politic, such officials serve to further normalize their more radicalized perspectives, particularly in the eyes of those previously latent extremists. Over time, it is no longer just the coded or explicit rhetoric that activates latent extremists within the broader public, but any political action by those officials.

To help illustrate our argument, Figure 1 presents the relationship between political appeals and activating latent racial extremists. An individual's belief in society's acceptance of various racial attitudes can vary from racial acceptance (left arrow on the figure) to racial exclusion (right arrow on the figure). This belief, in turn, affects the likelihood that latent extremists will act on their beliefs, with latent extremists being least likely to act, when they believe that their views are not acceptable by society, and most likely to act, when they belief that their views are widely accepted by society. Political elites, meanwhile, can make racially exclusive references using either multivocal, coded, or explicit appeals.

In each case, the type of appeal signals information about the degree of racial animus that is accepted by society. Multivocal appeals, for example, imply that racially exclusive policies are not socially acceptable. In this case, only a relatively small number of latent extremists will be activated. Coded appeals are more explicit than multivocal appeals, and signal a greater degree of socially acceptable racial animus. Compared to multivocal appeals, we would expect that explicit appeals will activate a larger number of latent extremists. Finally, even the presence of explicit appeals itself indicates a degree of social acceptance of overt racial exclusion.

The final piece of our argument is that, once extremists no longer feel that their views are marginalized by the broader society, any type of political rhetoric may function as a trigger for action on their extremist beliefs. Violent expression of extremist views may be activated by one of two mechanisms: the enabling mechanism and the lashing-out mechanism. The enabling mechanisms is activated in response to statements that may be interpreted as congruent with extremist views and, thus justifying expression and action upon these views. In contrast, the lashing-out mechanism is triggered by the statements of the elected officials who condemn extremism: given the legitimating effect of the first mechanism, individuals with extremist views may feel compelled to act in response to attempts to silence or condemn the views they now perceive as part of the mainstream.

Importantly, once the extremist audiences have been brought into the political mainstream, neither enabling nor triggering a lashing-out effect requires explicit or even coded anti-minority content. Once activated, extremists may be triggered by any type of rhetoric or activity by the representatives they follow. Much like even innocuous statements of elected officials are often interpreted as positive or negative through a partisan lens (Marks et al., 2018), once an elected official is cast as either a supporter or an opponent of the extremist ideology, all of their subsequent activity will be interpreted by extremist audiences through an extremist ideological lens. In summary, our theoretical framework leads us to expect a positive association between hate crimes and statements by all elected officials with large extremist audiences.

Research Design

To evaluate our hypothesis, we examine the relationship between statements by elected officials with large white nationalist audiences and the number of hate crimes in the US. Our dependent variable is a daily count of hate crimes, or crimes motivated by prejudice based on race, nationality, religion, gender, or sexual orientation, etc. Data on hate crimes were manually coded by three independent coders, using ProPublica's Hate Crime Index (2018) to help identify a sample of relevant cases.[3] For each case, we recorded the date and type of the incident (e.g., assault, vandalism); the type variable was later used to code a binary measure of whether or not an incident involved the threat or use of physical violence.

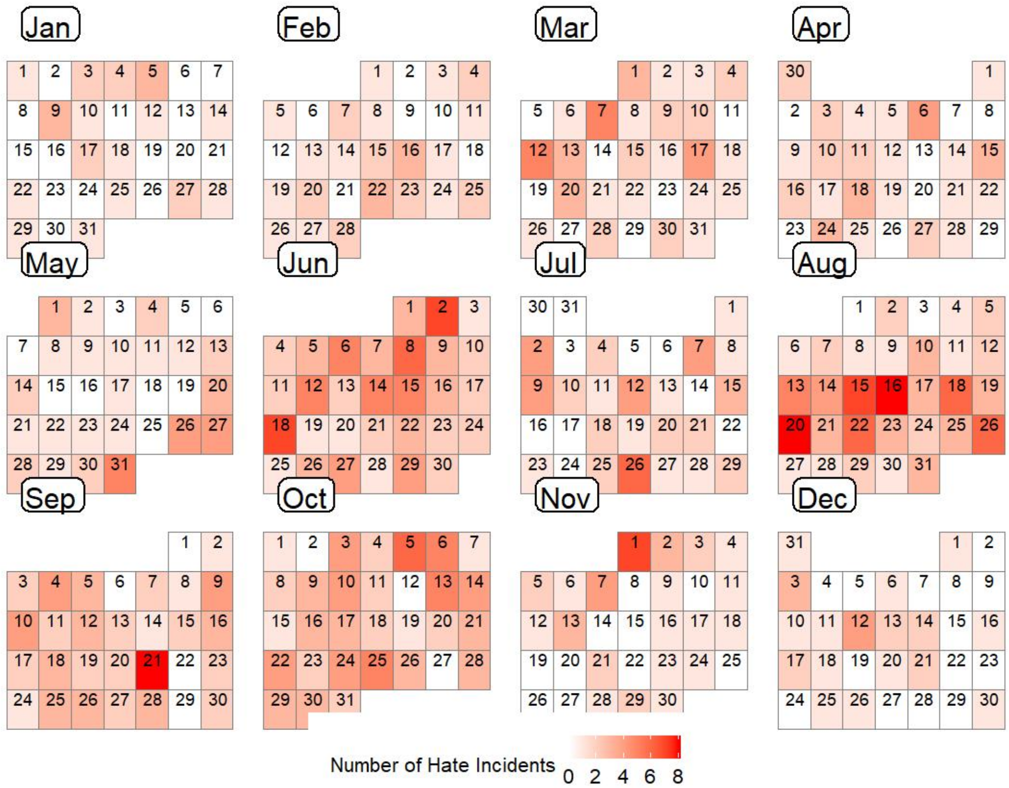
The final dataset contains information on 697 unique hate crimes between January 1, 2017–Apr 5, 2018. These include incidents ranging from racist graffiti, cross-burnings, and vandalism of mosques and synagogues to explicit threats, assaults, murder, and mass murder. For robustness, we also perform analysis on a subset of 313 incidents involving violence (e.g., assault, murder). Figure 2 shows an overview of the temporal distribution of hate incidents in our data for 2017. There is considerable day-to-day variation in the number of hate incidents, although the overall counts of incidents are larger during warmer months.

The primary independent variable—statements by elected officials with large white nationalist followings—is measured as the aggregated daily count of tweets for the US representatives.[4] The construction of this variable involved several steps. We started by downloading

[3]ProPublica provides a list of news reports of hate incidents that were published within our time period. We collected our data by manually reading each story within this list. We removed any duplicate stories. In cases of broken links, we performed additional searches using the article title.

[4]All Twitter data were obtained using the `rtweet` package in R(Kearney, 2018). While Twitter imposes

Figure 2: Hate Incidents in the US, 2017



all tweets for each representative that fall between February 15, 2017 to April 5, 2018. We use the official Twitter accounts of the members of the US House of Representatives for the 115th Congress.[5] The advantage of using social media to measure public statements is that this approach provides a rather exhaustive coverage (every US representative has a Twitter account and most tweet daily). Since the purpose of these accounts is to publicize member's activity, tweets often make references to public events, interviews, and other relevant actions of the representative. Using social media content also allows for the most direct test of our

some limitations on what can be harvested from their publicly available APIs, these limitations are most important for collecting data from the stream API. Downloading data on specific user accounts and their followers—the type of data used in this paper—involves using the rest API, which is a static archive of Twitter data. The only Twitter-imposed limitations on access to these data—access to only the last 2300 tweets for any account and the time limit of 75000 accounts per 15 minutes—did not affect our ability to obtain the data of interest.

[5]The Social Feed Manager project at the George Washington University Library has made a list of the Twitter accounts for the members of the house and Senate available at `https://gwu-libraries.github.io/sfm-ui/posts/2017-05-23-congress-seed-list`. We checked and amended this list as necessary to account for special elections or other changes.

theory, as social media outlets provide the fastest, most direct, and unmoderated channel of communication between the representatives and the public.

Next, we downloaded account information (Twitter handle, name, user bio) on the Twitter followers of each representative's account. In order to identify US representatives with the largest white nationalist following, we searched followers' bios for the key words commonly associated with white nationalism: white, alt-right, Aryan, ethno-nationalist, identitarian, and "14words".[6] Extracting information from the user bios has two advantages. First, it reduces the likelihood that our numbers are inflated by bots (machine-created accounts, usually intended to inflate a user's number of followers, likes, retweets, or to post automated comments), as bot accounts usually do not provide bios. Second, since Twitter limits user bio descriptions to 160 characters, we can assume that bios tend to reflect the most salient aspects of the user's identity. We validated the classification algorithm by reviewing a subset of accounts that were flagged as belonging to a white nationalist. We used the results of this search to construct a ratio of white nationalist followers to the total number of followers for each US representative.[7]

Table 1 provides a list of the 10 US Representatives with the the highest proportions of white nationalist followers. The list offers some face validity that our measure identifies US representatives whose activity is of most interest to white nationalist audiences. The first on the list is Lacy Clay, an African American representative from Missouri's first congressional district that includes the cities of St. Louis and Ferguson. Representative Clay is most likely to have attracted attention from white nationalist audiences when he filed a lawsuit to prevent the US Capitol building from removing a painting entitled "Racism Kills" (Hsu, 2017). Second and third on the list are Steve King and Steve Scalise, who are infamous for

[6]We also searched for alternative spellings and capitalization, e.g., altright, alt right, alt-right. After examining the data, we removed words and phrases to reduce the false positives rate, e.g., White House, teeth whitening, Go White, Red White and Blue, white silence is violence, black and white, white-label, whiteboard, whitetails, whitefield, white sox.

[7]Admittedly, our approach is likely to undercount the true proportion of white nationalist followers. Since the purpose of the ratio measure is to identify US representatives with largest white nationalist following, undercounting white nationalist followers will not affect our analysis, i.e. there is a strong correlation between our measure and the true measure.

Table 1: White Nationalist Twitter Following of US Representatives, Top 10.

	Name	District	White Nationalists
1	Lacy Clay	MO-1	12.50
2	Steve Scalise	LA-1	6.90
3	Steve King	IA-4	6.56
4	Andy Biggs	AZ-5	4.96
5	Matt Gaetz	FL-1	4.94
6	Francis Rooney	FL-19	4.63
7	Lee Zeldin	NY-1	4.54
8	Brendan Boyle	PA-13	4.36
9	John Ratcliffe	TX-4	4.31
10	Raja Krishnamoorthi	IL-8	4.25

Note: Number of white nationalist followers on Twitter (per 1,000 followers).

their racially charged rhetoric. For example, King once said that, "Western civilization is a superior civilization" compared to "Middle Eastern civilization" in an interview on Breitbart radio where he was defending a tweet decrying the fall of Western civilization, and on a separate talk radio appearance, King recommended a racist novel *Camp of Saints*, a story about the invasion of Europe by non-white immigrants (Gabriel, 2019). Steve Scalise once referred to himself as "David Duke without the baggage" (Grace, 2014).[8] Other prominent members of the list are Andy Biggs, who among other controversies, wrote a letter calling for the pardon of Joe Arpaio, a sheriff prosecuted for racially profiling minorities in defiance of a federal court order. Matt Gaetz, fifth on the list, invited Charles Johnson, an alt-right Holocaust denier, to attend the 2018 State of the Union.

Consistent with the two mechanisms developed in the theoretical section, white nationalist audiences follow two types of elected officials: those whose words or actions may be viewed as condoning white nationalist ideology, as well as minority elected officials who are outspoken about condemning white nationalism. In a political environment, in which elected officials refer to racially divisive issues in explicit rather than coded terms, activity (e.g. tweets) by both types of officials may elicit violent action against minorities: the former may enable expression of extremist views, while the latter may trigger the lashing-out effect.

[8]David Duke is a former leader of the white supremacist organization, the Ku Klux Klan.

Figure 3: White Nationalist Twitter Following of US Representatives

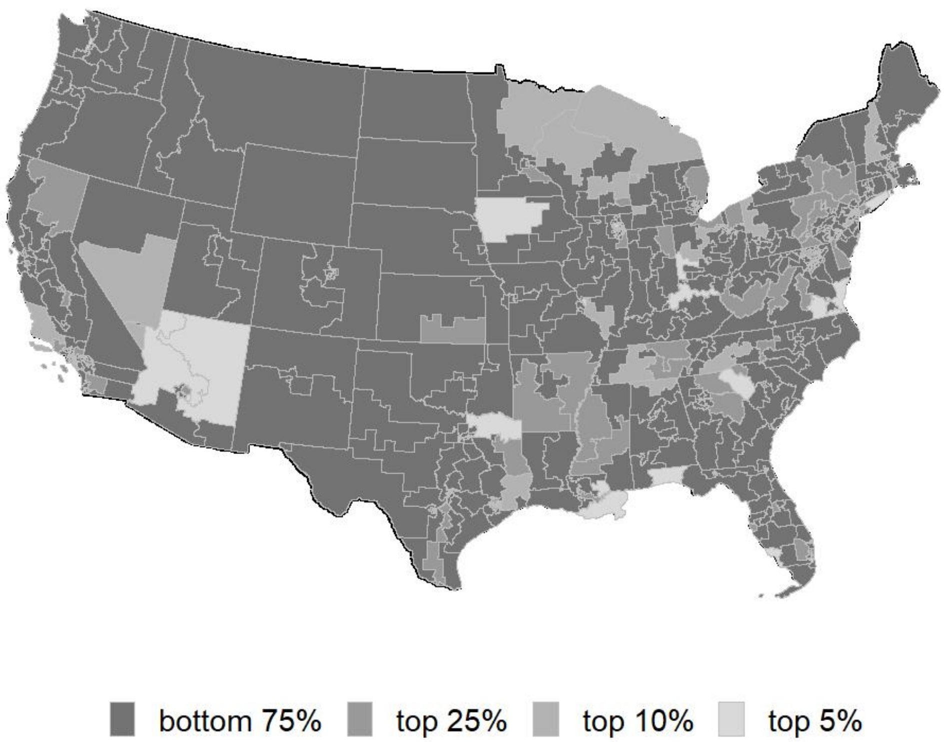


Figure 3 summarizes the white nationalist scores for all the members of the U.S. House of Representatives for the districts located in the continental US. The color intensity of the congressional districts corresponds to the proportion of whine nationalists per thousand followers. The darkest gray districts fall in the lowest 75% of districts. The shading becomes lighter as the white nationalist scores increase to the top 25%, 10%, and 5%. The spatial variation shown in the figure highlights that the size of white nationalist following is not a function of political party or ideology. Neither can we see any obvious regional patterns.

The final step in constructing our independent variable is to aggregate the daily count of tweets for the US representatives with the largest white nationalist followings. For the purposes of robustness, we create this measure for the top 25%, 10%, and 5% members of the House of Representatives. We log-normalize the final measure to obtain the variable *Log Tweets*.

We test our hypothesis by estimating a Poisson exponentially weighted moving average

(PEWMA) model (Brandt et al., 2000).[9] The PEWMA model has two components, the transition equation and the measurement equation. The transition equation describes the dynamic process that causes past events to affect current events (Brandt et al., 2000, 827).[10] The ω coefficient is the discount rate associated with this weighted mean, and varies between 0 and 1. An ω equal to 1 indicates that there is no movement in the mean, previously observed events have no effect on contemporaneous events and the PEWMA model is equivalent to a standard Poisson regression. If $\omega < 1$ the mean varies and the extent to which the mean varies increases as ω approaches 0. The measurement equation describes the process that generates the observed events. The measurement equation includes the covariates that are hypothesized to affect the mean number of events. The estimated coefficients on the covariates are interpreted the same as those from a standard Poisson model.

We also include a number of controls in our regression models. First, we include the daily counts of tweets by President Trump, and the daily presidential approval rating obtained from FiveThirtyEight (2019). Several prominent members of the white nationalist community in the United States have made public pronouncements about the significance of the President and his policies to the white nationalist movement. For example, David Duke, a former leader of the Klu Klux Klan (KKK), told a group of people at the Unite the Right rally in Charlottesville Virginia that, "We are determined to take our country back . . . We are going to fulfill the promise of Donald Trump. That's what we believed in. That's why we voted for Donald Trump, because he said he's going to take our country back" (quoted in Nelson 2017). To the extent that this kind of sentiment is shared by a large number of white nationalists, it is possible that increases in Donald Trump's approval ratings may embolden white nationalists to commit more hate crimes. Or, at least, that the series may covary with political conditions relevant to this activity such that the approval series may pick up some

[9]We chose to estimate a PEWMA model after evaluating the autocorrelation function (ACF) for persistence in the time series of our dependent variable Brandt et al. (2000) and Brandt and Williams (2001). The ACF for hate incidents time series is characterized by a slow decay which indicates a persistent process which is best modeled with a PEWMA model.

[10]The state-space equation specifies that the mean follows a random walk process.

of this variation.

We also control for daily economic and weather conditions, as well as other daily events. We proxy economic conditions using the daily change in the Dow Jones Industrial Average.[11] We measure daily changes in the national average temperature and precipitation, as crime rates are known to correlate with weather conditions (Field, 1992). We include a dummy variable for the Charlottesville rally, as well as an indicator the following week after the rally, to determine whether they motivated additional violence. Finally, we include an indicator for weekends, as the rate of hate crimes should increase on days when people are less frequently at work. The results are presented in the next section.

Results

The results from our analysis are presented in Table 2. Models 1-3 are estimated on all hate incidents, while Models 4-6 are estimated on the subset of only violent incidents. Models 1 and 4 use the log of the aggregated daily tweets for the representatives whose white nationalist following ranks in the top 25%. Models 2 and 5 use the top 10% and models 3 and 6 use the aggregate daily tweets for the representatives whose white nationalist following ranks in the top 5%. The bottom panel of Table 2 shows the sample size (N), the Akaike information criteria (AIC), and a Wald test for the null hypothesis that the ω parameter is equal to one ($\text{Wald}_{\hat{\omega}}$). The ω parameter is the discount rate from the state-space equation. If $\omega = 1$, the PEWMA model reduces to a poisson model. The value of about 0.9 on this parameter indicates moderate persistence in the hate incidents time-series. This result is intuitive. To the extent that news coverage of hate crimes could motivate similar crimes we would expect that this kind of activity would occur within a short time frame.

The results presented in Table 2 are consistent with our hypothesis. The coefficients on *Log Tweets* are positive and statistically significant in every model. The rate of hate incidents

[11] Since markets are closed on weekends and major holidays, we impute the missing values using a Kalman filter.

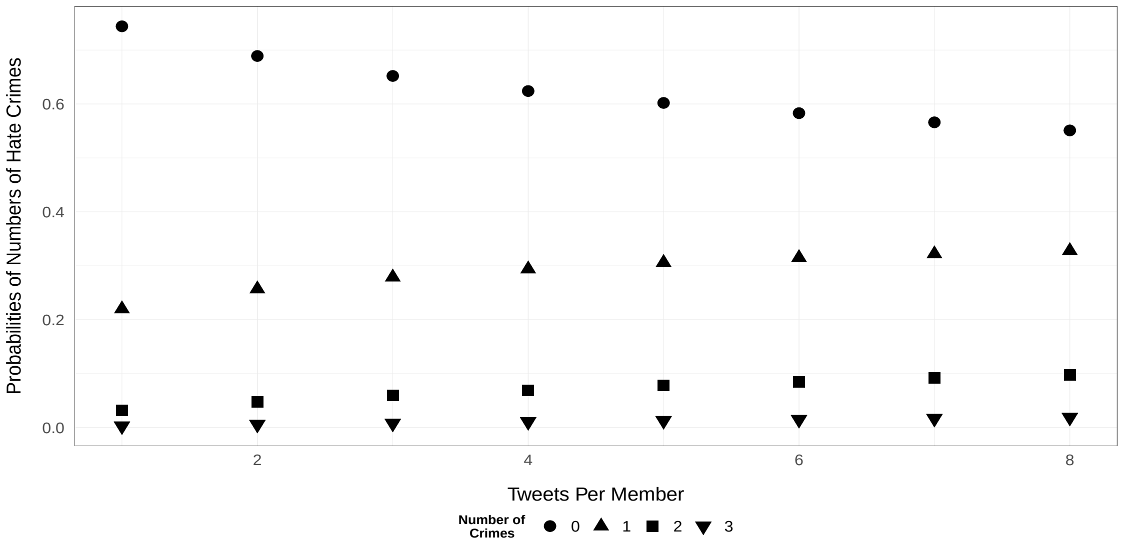
Table 2: PEWMA Models of Twitter activity and Hate Crimes

	All Hate Crimes			*Violent Hate Crimes*		
	Top 25%	Top 10%	Top 5%	Top 25%	Top 10%	Top 5%
Log Tweets	0.338***	0.313***	0.294***	0.360***	0.282**	0.287**
	(0.111)	(0.114)	(0.108)	(0.133)	(0.136)	(0.129)
Trump Tweets	0.016*	0.015^{+}	0.015^{+}	0.009	0.009	0.008
	(0.009)	(0.009)	(0.009)	(0.012)	(0.012)	(0.012)
Approval	-0.072	-0.066	-0.065	-0.068	-0.057	-0.056
	(0.068)	(0.068)	(0.068)	(0.067)	(0.068)	(0.069)
Dow Jones	8.038	8.609	9.215	9.052	9.977	10.591
	(9.156)	(9.114)	(9.127)	(11.307)	(11.297)	(11.295)
Charlottesville	-1.200*	-1.245**	-1.199*	-0.771	-0.808	-0.770
	(0.617)	(0.617)	(0.616)	(0.636)	(0.637)	(0.636)
Charlottesville Week	0.674***	0.704***	0.707***	0.644**	0.669**	0.673**
	(0.257)	(0.256)	(0.256)	(0.296)	(0.298)	(0.299)
Weekends	0.392**	0.347**	0.264**	0.473**	0.368**	0.314**
	(0.158)	(0.156)	(0.133)	(0.189)	(0.187)	(0.160)
Temperature	0.118**	0.123**	0.124**	0.158**	0.165***	0.166***
	(0.050)	(0.050)	(0.051)	(0.062)	(0.062)	(0.062)
Precipitation	-0.025	-0.023	-0.026	-0.004	-0.001	-0.004
	(0.038)	(0.038)	(0.038)	(0.046)	(0.046)	(0.046)
$\hat{\omega}$	0.897***	0.896***	0.894***	0.935***	0.931***	0.929***
	(0.023)	(0.022)	(0.023)	(0.019)	(0.019)	(0.019)
AIC	1295.309	1296.963	1296.999	1083.448	1086.383	1085.708
$\text{Wald}_{\hat{\omega}}$	20.401***	21.379***	21.948***	11.516***	12.487***	13.409***
N	411	411	411	411	411	411

Note: $^{***}p < .01$, $^{**}p < .05$, $^{*}p < .1$ (two tail); $^{+}p < .1$ (one tail). The columns show the coefficients for the Poisson exponentially weighted moving average (PEWMA) regressions of hate crimes and violent hate crimes. The standard errors are shown in parentheses. The regressions use the logged aggregated tweet counts from the top 25%, 10%, and 5% of US representatives in terms of identifiable white nationalist followers as a proportion of total followers. The $\hat{\omega}$ coefficient is the estimated discount rate for past observations. The $\text{Wald}_{\hat{\omega}}$ test is a test of the null that $\hat{\omega} = 1$. If $\hat{\omega} = 1$, the PEWMA model reduces to a standard Poisson model.

increases in response to the Twitter activity of the US representatives with the largest white nationalist followings. Figure 4 shows the predicted probabilities for different counts of hate incidents as a function of these representatives' Twitter activity. These probabilities are calculated using coefficients in Model 1 of Table 2, setting the control variables to their mean

Figure 4: Predicted Probabilities of Observing Daily Hate Crimes by Number of Tweets per Member (Top 25%)



Note: The y-axis shows the probability of observing N hate crimes per day where N varies from 0 hate crimes to 3 hate crimes. The x-axis shows the number of tweets per US representative for the top 25% of US representatives in terms of proportion of followers that identify as white nationalists. There are 107 members in the top 25%, so 1 Tweet is 107 total tweets in a day. The probabilities are based on Model 1 from Table 2.

and modal values, as applicable. We can see that the effect of rhetoric on hate incidents is not only statistically but also substantively significant. As the number of tweets per representative increases from 1 to 8, the probability of observing no hate incidents rapidly decreases from 0.75 to 0.55—an almost 30 percent decrease. We also see that a change from 1 to 8 tweets per representative is associated with a 50 percent increase in the probability of observing 1 incident (from about 0.22 to 0.33), and an 150 percent increase of in the probability of observing 2 incidents (from 0.04 to 0.1).

Among the control variables, the coefficient on *Trump Tweets* is positive and statistically significant in the models that include all hate incidents, albeit at the 0.1 (two-tailed) and 0.1 (one-tailed) thresholds. This suggests that tweets by the President may incite hate incidents, especially at the low intensity level (e.g. racist graffiti). The coefficients for *Charlottesville* and *Charlottesville Week* are both statistically significant but the effect on *Charlottesville* is negative. There were fewer hate crimes committed on the day of the Unite the Right rally, but the media coverage of the events in Charlottesville appears to have activated white nationalists across the country, increasing the rate of hate crimes. Also consistent with expectations, the rate of hate incidents increases on weekends and is also positively related to increases in *Temperature*.

Robustness Checks

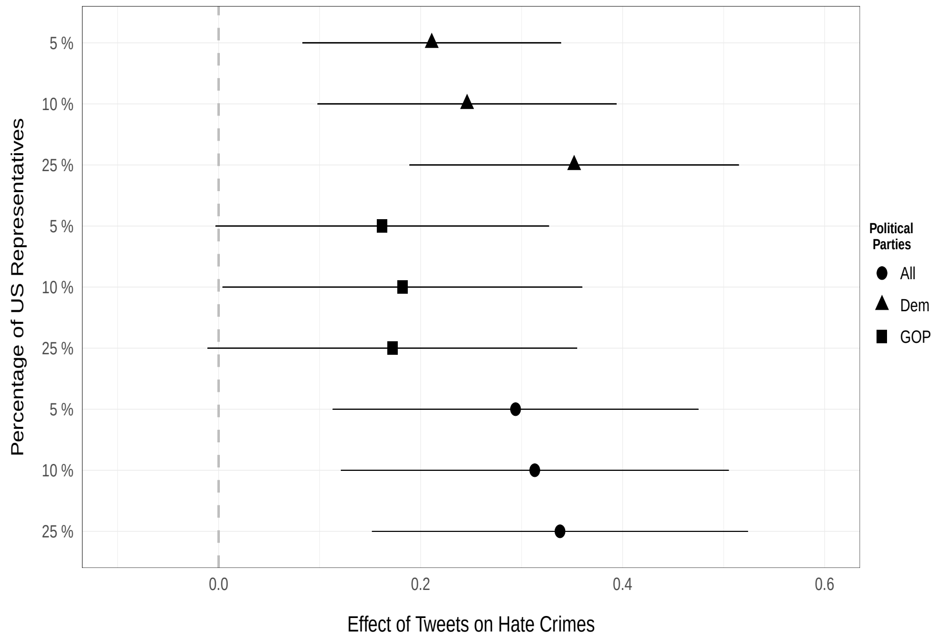
We conduct a series of robustness checks, the detailed results of which are presented in the Online Appendix. First, we further subset the *Log Tweets* variable by political party to assess if one of the parties is driving the results.[12] These results are presented in Figure 5. These results show that the effect of rhetoric on hate incidents is not easily attributable to one or the other party: the coefficients on each are positive and statistically significant, and are very close in size.

We also perform a sensitivity analysis of our *Log Tweets* measure to ensure that the results are not disproportionately influenced by a small number of representatives. For this, we estimate 107 models (there are 107 representatives in the top 25th percentile in terms of white nationalist followings), each with the *Log Tweets* variable coded omitting one of the representatives at a time. We then compare the coefficients in each of these models to those in the main model. The comparisons reveal no evidence that any of the individual representatives significantly alter the results.

Finally, we asked two research assistants to hand code tweets by the representatives whose white nationalist audiences rank in the top 10th percentile for explicit anti-minority content. We then re-estimated the model both replacing *Log Tweets* measure with this variable, and including both variables. The coefficient on *Explicit Tweets* is not statistically significant in any of the models, while the coefficient on *Log Tweets* is unaffected in size or significance. This result is consistent with our theoretical argument that, once an elected official expresses an opinion that validates extremism, the entirety of his or her activity adds

[12]We estimate the effect of tweets by each party in separate model, as the number of tweets is correlated at approximately 0.7.

Figure 5: PEWMA Coefficients for Different Political Parities



Note: PEWMA coefficients are plotted with 90% Confidence intervals against different subsets of the sample. The *GOP* coefficients are taken from models that only have Republican representatives, the *Dem* coefficients are taken form models that only have Democratic representatives, and the *All* coefficients are taken from models that have the full sample, like the models presented in Table 2. The full models are presented in the appendix.

additional validation to these views, at least in the eyes of the extremist audiences (as the saying goes, "if you look hard enough for something, you will find it"). This result may also indicate that, while there are some prominent examples of explicit white nationalist rhetoric, most of such rhetoric remains coded or at least ambiguous enough to evade being flagged by research assistants.

Conclusion

Political rhetoric affects political behavior. We argue that this extends to political violence. Elite political rhetoric conditions what individuals believe regarding society's racial and ethnic attitudes. Dog whistles position elites as supporting or condoning radical ideologies. For the acolytes of these ideologies, the subsequent activities of these elites become affirmations of their beliefs. We hypothesize that these activities are positively correlated with racially-motivated crimes. We test this hypothesis using data collected from Twitter. We develop a novel means of identifying radicalized constituencies within the Twitter sphere. We find that the Twitter activity of house members with the largest radicalized audiences is related to violence against minorities. These results have important theoretical and normative implications.

First, our results show that a failure to categorically reject radical ideology has serious consequences. The rhetorical choices politicians makes matter. We are not simply arguing that hate speech by politicians causes hate crimes. The connection is more pernicious. Once white nationalists perceive an elected official as sympathetic to white nationalism, all of the actions of the official become an affirmation of the ideology. These affirmations empower and embolden members of the white nationalist community. Once extremist views receive validation by some elected officials, condemning these views may also elicit a lashing-out effect. We provide systematic evidence for our theoretical model. There is a robust relationship between the Twitter activity of house members— both those that have condoned and those that condemned white nationalist ideology—and hate crimes. Our auxiliary analyses show that it is not just the Tweets associated with racist organizations and policies that have an effect. Once elected officials activate an extremist audience by taking an explicit stance on the issue, all of their subsequent Twitter activity matters. Politicians cannot cordon off their dog whistle politics from their mainstream politics.

The finding that dog-whistle politics affects political violence raises a series of important questions that should be considered in future research. Politicians can attach themselves to radical ideologies with dog whistles, is there a way to disconnect? There is anecdotal evidence that radicalized constituencies exert pressure on politicians. It is possible that distancing oneself from these groups, once they have been emboldened, has political costs. This envisages a feedback loop where politicians have incentives to engage with hate groups, the engagement creates a growing atmosphere of hostility and violence, and the atmosphere creates additional incentives to continue engaging with extremist groups. These dynamics require further scrutiny.

Second, our argument about rhetorical choices and institutions can be tested in other contexts. When an elected official appeals to a radical ideology, the official's station as a member of government confers legitimacy to the ideology and the official's actions are a representation of the ideology. The official makes the ideology official. Consider the Indian prime minister Narendra Modi. Since his election in 2014, Modi has made explicit appeals to Hindu nationalist movements in India and pursued what has been described as a *Hindu first* political agenda. Modi rarely makes explicitly anti-Muslim statements, but his approach has caused Hindu nationalist views, and the corresponding anti-Muslim sentiment to "hit a higher crest than ever before" (Gettleman et al., 2019). Our theory and results explain this phenomenon. When Modi gives a speech about economic policy, the speech is not sundered from his status as a Hindu-nationalist figure, even if the speech does not contain Hindu-nationalist content. For Hindu nationalists, Modi is not the Indian prime minister giving a speech, he is the Hindu-nationalist prime minister giving a speech. When he makes identity based appeals to get elected, those appeals become a defining feature of his politics. Twitter engagement has been important in this context as well. Modi has been criticized for following Hindu nationalists on Twitter because it is seen as an endorsement of Hindu nationalist politics (Gettleman, 2017). We have seen these dynamics play out in the US and India, and we are starting to see it in parts of Europe and South America. Future research should endeavor to compare these dynamics across countries and ideologies.

Finally, our approach to studying Twitter activity offers a novel framework for scholars interested in studying public engagement and political behavior. Twitter is a distinct setting for political engagement. It is one of many ways politicians can connect with national audiences but it is one that we can observe in its entirety. Twitter timelines are public, as are the public profiles of the people that follow politicians on Twitter. This study is the first to use information about the way people present themselves on Twitter to understand the dynamics of political engagement with a distinct audience. We focus on white nationalism but one could also focus on more benign features of identity like sex, partisanship, or race. We think our efforts here represent a first step to a broader conversation about how this information can be harnessed.

In todays socio-political milieu, the words politicians use affect they way people view their actions. When politicians makes white nationalist appeals, attempt to strike a tone of conciliation with white nationalist views, or support policies supported by white nationalists; they tie themselves to white nationalism. Politicians seem to believe they can wear multiple hats, making different kinds of appeals to different groups to cobble together a base of support. While this might be true for policy issues like trade and abortion, it is not true for white nationalism. That hat doesn't come off. There are serious consequences to flirting with this ideology. Politicians should avoid indulging in dog-whistle politics and the public should hold those that do accountable.

References

Albertson, Bethany L. 2015. "Dog-whistle Politics: Multivocal Communication and Religious Appeals." *Political Behavior* 37(1):3–26.

Asal, Victor and Andrew Vitek. 2018. "Sometimes They Mean What They Say: Understanding Violence among Domestic Extremists." *Dynamics of Asymmetric Conflict* `https://doi.org/10.1080/17467586.2018.1470659`. Forthcoming.

Avdan, Nazli and Clayton Webb. 2019. "Not In My Back Yard: Public Perceptions and Terrorism." *Political Research Quarterly* 72(1):90–103.

Bail, Christopher A. 2012. "The Fringe Effect: Civil Society Organizations and the Evolution of Media Discourse about Islam since the September 11th Attacks." *American Sociological Review* 77(6):855–879.

Barceló, Joan and Elena Labzina. 2018. "Do Islamic State's Deadly Attacks Disengage, Deter, or Mobilize Supporters?" *British Journal of Political Science* `https://www.researchgate.net/publication/325054772_Do_Islamic_State's_Deadly_Attacks_Disengage_Deter_or_Mobilize_Supporters`. Forthcoming.

Barkun, Michael. 2017. "President Trump and the "Fringe"." *Terrorism and Political Violence* 29(3):437–443.

Bartels, Larry M. 2018. "Partisanship in the Trump Era." *Journal of Politics* 80(4):1483–1494.

Boucher, Jean-Christophe and Cameron G. Thies. 2019. ""I Am a Tariff Man": The Power of Populist Foreign Policy Rhetoric under President Trump." *Journal of Politics* 81(2):712–722.

Brandt, Patrick T. and John T. Williams. 2001. "A Linear Poisson Autoregressive Model: The Poisson AR(p) Model." *Political Analysis* 9(2):164–184.

Brandt, Patrick T., John T. Williams, Benjamin O. Fordham and Brain Pollins. 2000. "Dynamic Modeling for Persistent Event-count Time Series." *American Journal of Political Science* pp. 823–843.

Brewer, Marilynn B. 1999. "The Psychology of Prejudice: Ingroup Love and Outgroup Hate?" *Journal of Social Issues* 55(3):429–444.

Brewer, Marilynn B. 2007. "The Importance of Being We: Human Nature and Intergroup Relations." *American Psychologist* 62(8):726–738.

Canetti-Nisim, Daphna, Eran Halperin, Keren Sharvit and Stevan E Hobfoll. 2009. "A New Stress-based Model of Political Extremism: Personal Exposure to Terrorism, Psychological Distress, and Exclusionist Political Attitudes." *Journal of Conflict Resolution* 53(3):363–389.

Chan, Jason, Anindya Ghose and Robert Seamans. 2015. "The Internet and Racial Hate Crime: Offline Spillovers from Online Access." *MIS Quarterly* 40(2):381–403.

Ditonto, Tessa M., Richard R. Lau and David O. Sears. 2013. "AMPing Racial Attitudes: Comparing the Power of Explicit and Implicit Racism Measures in 2008." *Political Psychology* 34(4):487–510.

Fearon, James D. and David D. Laitin. 2000. "Violence and the Social Construction of Ethnic Identity." *International Organization* 54(4):845–877.

Feldman, Stanley and Leonie Huddy. 2005. "Racial Resentment and White Opposition to Race-conscious Programs: Principles or Prejudice?" *American Journal of Political Science* 49(1):168–183.

Field, Simon. 1992. "The Effect of Temperature on Crime." *The British Journal of Criminology* 32(3):340–351.

FiveThirtyEight. 2019. "Trump Approval Ratings.".
URL: *https://projects.fivethirtyeight.com/trump-approval-ratings/*

Gabriel, Trip. 2019. "A Timeline of Steve King's Racist Remarks and Divisive Actions." *The New York Times* .
URL: *https://www.nytimes.com/2019/01/15/us/politics/steve-king-offensive-quotes.html*

Garcia, Blake E. and Cameron Wimpy. 2016. "Does Information Lead to Emulation? Spatial Dependence in Anti-government Violence." *Political Science Research and Methods* 4(1):27–46.

Gettleman, Jeffrey. 2017. "India's Modi Criticized for Following Twitter Feed Tied to Nasty Post." *The New York Times* September, 8.
URL: *https://www.nytimes.com/2017/09/08/world/asia/india-modi-twitter.html*

Gettleman, Jeffrey, Kai Shultz, Suhasini Raj and Hari Kumar. 2019. "Under Modi, a Hindu Natinonalist Surge Has Further Divided India." *The New York Times* April 11.
URL: *https://www.nytimes.com/2019/04/11/world/asia/modi-india-elections.html*

Gibler, Douglas M., Marc L. Hutchison and Steven V. Miller. 2012. "Individual Identity Attachments and International Conflict: The Importance of Territorial Threat." *Comparative Political Studies* 45(12):1655–1683.

Glaser, Jack, Jay Dixit and Donald P. Green. 2002. "Studying Hate Crime with the Internet: What Makes Racists Advocate Racial Violence?" *Journal of Social Issues* 58(1):177–193.

Grace, Stephanie. 2014. "Stephanie Grace: Scalise's pitch to Duke supporters seems plausible." *The Advocate* December, 31.
URL: $https://www.theadvocate.com/baton_rouge/opinion/stephanie_grace/article_5d6ad63a-caed-5ad7-a377-c5cb0d41d1ac.html$

Green, Donald P., Robert P. Abelson and Margaret Garnett. 1999. The Distinctive Political Views of Hate-crime Perpetrators and White Supremacists. In *Cultural Divides: Understanding and Overcoming Group Conflict*, ed. Deborah Prentice and Dale Miller. New York: Russell Sage Foundation.

Gurr, Ted Robert. 2000. *Peoples Versus States: Minorities at Risk in the New Century*. US Institute of Peace Press.

Hate Crime Index. 2018. Collaboration between Goolge News Lab and Propublica. Available at `https://projects.propublica.org/hate-news-index/`.

Hsu, Spencer S. 2017. "Congressman Sues over Removal of Ferguson-inspired Student Painting from U.S. Capitol." *The Washington Post* February 21.
URL: *https://www.washingtonpost.com/local/public-safety/congressman-sues-over-removal-of-ferguson-inspired-student-painting-from-us-capitol/2017/02/21/57fbcd0c-f851-11e6-bf01-d47f8cf9b643*$_s$*tory.html?noredirect = onutm*$_t$*erm* = .931a3a700438

Hughey, Matthew W. and Gregory S. Parks. 2014. *The Wrongs of the Right: Language, Race, and the Republican Party in the Age of Obama*. NYU Press.

Hurwitz, Jon and Mark Peffley. 2005. "Playing the Race Card in the Post–Willie Horton Era: The Impact of Racialized Code Words on Support for Punitive Crime Policy." *Public Opinion Quarterly* 69(1):99–112.

Iyengar, Shanto and Donald R. Kinder. 2010. *News that Matters: Television and American Opinion*. University of Chicago Press.

Iyengar, Shanto and Kyu S. Hahn. 2009. "Red Media, Blue Media: Evidence of Ideological Selectivity in Media Use." *Journal of Communication* 59(1):19–39.

Jacobson, Gary C. 2018. "The Effects of the Early Trump Presidency on Public Attitudes Toward the Republican Party." *Presidential Studies Quarterly* 48(3):404–435.

Jamieson, Kathleen Hall and Doron Taussig. 2017. "Disruption, Demonization, Deliverance, and Norm Destruction: The Rhetorical Signature of Donald J. Trump." *Political Science Quarterly* 132(4):619–650.

Jungkunz, Sebastian, Marc Helbling and Carsten Schwemmer. 2019. "Xenophobia Before and After the Paris 2015 Attacks: Evidence from a Natural Experiment." *Ethnicities* 19(2):271–291.

Kalmoe, Nathan P. 2014. "Fueling the Fire: Violent Metaphors, Trait Aggression, and Support for Political Violence." *Political Communication* 31(4):545–563.

Kearney, Michael W. 2018. "rtweet: Collecting Twitter data." *Comprehensive R Archive Network* .

Kinder, Donald R. and Cindy D. Kam. 2010. *Us Against Them: Ethnocentric Foundations of American Opinion*. University of Chicago Press.

Kinder, Donald R. and David O. Sears. 1981. "Prejudice and Politics: Symbolic Racism Versus Racial Threats to the Good Life." *Journal of Personality and Social Psychology* 40(3):414–431.

Kinder, Donald R. and Lynn M. Sanders. 1996. *Divided by Color: Racial Politics and Democratic Ideals*. University of Chicago Press.

King, Gary, Benjamin Schneer and Ariel White. 2017. "How the News Media Activate Public Expression and Influence National Agendas." *Science* 358(6364):776–780.

King, Gary, Jennifer Pan and Margaret E. Roberts. 2017. "How the Chinese Government Fabricates Social Media Posts for Strategic Distraction, Not Engaged Argument." *American Political Science Review* 111(3):484–501.

López, Ian Haney. 2015. *Dog Whistle Politics: How Coded Racial Appeals have Reinvented Racism and Wrecked the Middle Class*. Oxford University Press.

Mansfield, Edward D. and Jack Snyder. 2002. "Democratic Transitions, Institutional Strength, and War." *International Organization* 56(2):297–337.

Marks, Joseph, Eloise Copland, Eleanor Loh, Cass R. Sunstein and Tali Sharot. 2018. "Epistemic Spillovers: Learning Others' Political Views Reduces the Ability to Assess and Use Their Expertise in Nonpolitical Domains." *Cognition* 188(1):74–84.

Mendelberg, Tali. 2001. *The Race Card: Campaign Strategy, Implicit Messages, and the Norm of Equality*. Princeton University Press.

Miller, Steven V. 2017. "Economic Threats or Societal Turmoil? Understanding Preferences for Authoritarian Political Systems." *Political Behavior* 39(2):457–478.

Mitchell, Sara M. and Brandon C. Prins. 2004. "Rivalry and Diversionary Uses of Force." *Journal of Conflict Resolution* 48(937–61).

Mousseau, Demet Yalcin. 2001. "Democratizing with ethnic divisions: a source of conflict?" *Journal of Peace Research* 38(5):547–567.

Mozur, Paul. 2018. "A Genocide Incited on Facebook, With Posts from Myanmar's Military." *The New York Times* October 15.

Munger, Kevin, Richard Bonneau, Jonathan Nagler and Joshua A. Tucker. 2018. "Elites Tweet to Get Feet Off the Streets: Measuring Regime Social Media Strategies During Protest." *Political Science Research and Methods* `doi:10.1017/psrm.2018.3`. Forthcoming.

Nelson, Libby. 2017. ""Why We Voted for Donald Trump": David Duke Explains the White Supremacist Charlottesville Protests." *Vox* .
URL: *https://www.vox.com/2017/8/12/16138358/charlottesville-protests-david-duke-kkk*

Norton, Michael I. and Samuel R. Sommers. 2011. "Whites See Racism as a Zero-sum Game that They are Now Losing." *Perspectives on Psychological Science* 6(3):215–218.

Padgett, John F. and Christopher K. Ansell. 1993. "Robust Action and the Rise of the Medici, 1400–1434." *American journal of sociology* 98(6):1259–1319.

Panetta, Grace. 2019. "Twitter Reportedly Won't Use an Algorithm to Crack Down on White Supremacists because Some GOP Politicians Could End Up Getting Barred Too." *Business Insider* April 25.

Penn, Elizabeth Maggie. 2008. "Citizenship Versus Ethnicity: The Role of Institutions in Shaping Identity Choice." *Journal of Politics* 70(4):956–973.

Phillips, Brian J. 2017. "Deadlier in the U.S.? On Lone Wolves, Terrorist Groups, and Attack Lethality, Terrorism and Political Violence." *Terrorism and Political Violence* 29(3):533–549.

Phillips, Whitney. 2018. *The Oxygen of Amplification*. New York: Data & Society Research Institute.

Pickering, Jeffrey and Emizet F. Kisangani. 2005. "Democracy and Diversionary Military Intervention: Reassessing Regime Type and the Diversionary Hypothesis." *International Studies Quarterly* 49(1):23–43.

Quillian, Lincoln. 1995. "Prejudice as a Response to Perceived Group Threat: Population Composition and Anti-immigrant and Racial Prejudice in Europe." *American Sociological Review* 60(4):586–611.

Sears, David O., Carl P. Hensler and Leslie K. Speer. 1979. "Whites' Opposition to 'Busing': Self-interest or Symbolic Politics." *American Political Science Review* 73(2):369–84.

Sears, David O. and Donald R. Kinder. 1971. Racial Tensions and Voting in Los Angeles. In *Los Angeles: Viability and Prospects for Metropolitan Leadership*, ed. Werner Zvi Hirsch. New York: Praeger.

Sears, David O. and Patrick J. Henry. 2005. "Over Thirty Years Later: A Contemporary Look at Symbolic Racism." *Advances in Experimental Social Psychology* 37:95–150.

Sears, David O., Richard R. Lau, Tom R. Tyler and Harris M. Allen, Jr. 1980. "Self-interest Vs. Symbolic Politics in Policy Attitudes and Presidential Voting." *American Political Science Review* 74(3):670–684.

Tajfel, Henri. 1981. *Human Groups and Social Categories: Studies in Social Psychology*. Cambridge University Press.

Taub, Amanda and Max Fisher. 2018. "Where Countries Are Tinderboxes and Facebook Is a Match." *The New York Times* April 21.

Tilly, Charles. 2003. *The Politics of Collective Violence*. Cambridge University Press.

Tir, Jaroslav and Michael Jasinski. 2008. "Domestic-level Diversionary Theory of War: Targeting Ethnic Minorities." *Journal of Politics* 52(5):641–664.

Valentino, Nicholas A., Fabian G. Neuner and L. Matthew Vandenbroek. 2018. "The Changing Norms of Racial Political Rhetoric and the End of Racial Priming." *Journal of Politics* 80(3):757–771.

Valentino, Nicholas A., Vincent L. Hutchings and Ismail K. White. 2002. "Cues that Matter: How Political Ads Prime Racial Attitudes during Campaigns." *American Political Science Review* 96(1):75–90.

Ward, Michael D. and Kristian S. Gleditsch. 1998. "Democratizing for Peace." *American Political Science Review* 92(1):51–61.

Warren, T. Camber. 2015. "Explosive Connections? Mass Media, Social Media, and the Geography of Collective Violence in African States." *Journal of Peace Research* 52(3):297–311.

White, Ismail K. 2007. "When Race Matters and When It Doesn't: Racial Group Differences in Response to Racial Cues." *American Political Science Review* 101(2):339–354.

Wood, Elisabeth Jean. 2008. "The Social Processes of Civil War: The Wartime Transformation of Social Networks." *Annual Review of Political Science* 11:539–561.

Zaller, John R. 1992. *The Nature and Origins of Mass Opinion*. Cambridge university press.

Appendix for
The Effects of Dog Whistle Politics on Domestic Violence

For online publication only

Contents

1 Political Parties

This section demonstrates that the results presented in the body of the manuscript are not a consequence of the Twitter activity of any single party. Given the current political climate in the US, it is natural to question whether the Twitter activity of house members from both of the major political parties could be related to hate crimes. It is possible that a majority of the covariation between hate crimes and Twitter activity reported in Table 2 can be attributed to Republicans. After all, the highest profile members of the house that have been associated with white nationalists, Steve King (Gabriel 2019) and Steve Scalise (Costa and O'Keefe 2014), are members of the Republican party and it is Republicans who have been accused by Federal judges of targeting African American voters with "surgical precision" (Ingraham 2016).

We tested the conjecture that Republicans were driving the results by subsetting the sample by political party. The results from these models are presented in Figure 5 of the manuscript, and are taken from the the models presented here in Table A.1. All the coefficients are positive and the overlapping confidence intervals for the coefficients suggest that the effects cannot be distinguished from one another. These results militate against the interpretation that one party is driving the results.

The first column of Table A.1 lists the independent variables. The coefficients for the Twitter timelines (Log Tweets) are the coefficients that are presented in Figure 5 of the manuscript. The remaining variables are included as controls. Like the tables presented in the paper, the $\hat{\omega}$ coefficient reflects the discounting of previous observations of the mean count estimate, with values of $\hat{\omega}$ closer to 1 reflecting more discounting. The bottom panel of the table presents the Akaike Information Criteria (AIC), sample sizes (N), and Wald test that $\hat{\omega} = 1$. The first three models, presented in columns 1 through 3, use the Twitter timelines of Republican house members. Columns 4 through 6 present the results from the models for the Democratic house members. The percentages are referencing the top 25%, 10%, and 5% of members from each party in terms of white nationalist followers.

Table A.1: PEWMA Models of Twitter activity and Hate Crimes by Political Party

	Republicans			*Democrats*		
	Top 25%	Top 10%	Top 5%	Top 25%	Top 10%	Top 5%
Log Tweets	0.172+	0.182*	0.162+	0.352***	0.246***	0.211***
	(0.109)	(0.106)	(0.098)	(0.097)	(0.088)	(0.076)
Trump Tweets	0.017*	0.017*	0.017*	0.015+	0.015+	0.015+
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Approval	-0.058	-0.061	-0.064	-0.083	-0.065	-0.058
	(0.068)	(0.068)	(0.068)	(0.068)	(0.068)	(0.068)
Dow Jones	8.931	9.409	9.689	8.054	8.694	9.67
	(9.132)	(9.113)	(9.154)	(9.216)	(9.207)	(9.207)
Charlottesville	-1.251**	-1.28**	-1.259**	-1.14*	-1.174	-1.078*
	(0.62)	(0.619)	(0.618)	(0.616)	(0.616)	(0.616)
Charlottesville Week	0.779***	0.787***	0.775***	0.578**	0.639*	0.628**
	(0.26)	(0.259)	(0.258)	(0.259)	(0.257)	(0.259)
Weekends	0.200	0.176	0.127**	0.408***	0.309**	0.224*
	(0.159)	(0.139)	(0.121)	(0.144)	(0.144)	(0.121)
Temperature	0.124**	0.124**	0.128**	0.119**	0.127**	0.123**
	(0.051)	(0.051)	(0.051)	(0.050)	(0.050)	(0.051)
Precipitation	-0.026	-0.027	-0.028	-0.020	-0.016	-0.016
	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)
$\hat{\omega}$	0.895***	0.895***	0.894***	0.897***	0.896***	0.894***
	(0.023)	(0.022)	(0.022)	(0.023)	(0.023)	(0.023)
AIC	1302.054	1301.617	1301.879	1291.155	1296.72	1296.826
$\text{Wald}_{\hat{\omega}}$	21.478***	21.953***	22.700***	20.178***	21.103***	20.825***
N	411	411	411	411	411	411

Note: *** $p < .01$, ** $p < .05$, * $p < .1$ (two tail); + $p < .1$ (one tail). The columns show the coefficients for the Poisson exponentially weighted moving average (PEWMA) regressions of hate crimes and violent hate crimes. The standard errors are shown in parentheses. The regressions use the logged Twitter timelines from the top 25%, 10%, and 5% of US representatives in terms of identifiable white nationalist followers as a proportion of total followers. The $\hat{\omega}$ coefficient is the estimated discount rate for past observations. The $\text{Wald}_{\hat{\omega}}$ test is a test of the null that $\hat{\omega} = 1$. If $\hat{\omega} = 1$, the PEWMA model reduces to a standard Poisson model.

The results presented in Table A.1 are consistent with the results presented in the body of the paper. The significance level of some of the coefficients change but none of the inferences from the auxiliary models are different from the inferences from the results presented in Table 2.

2 Legislators

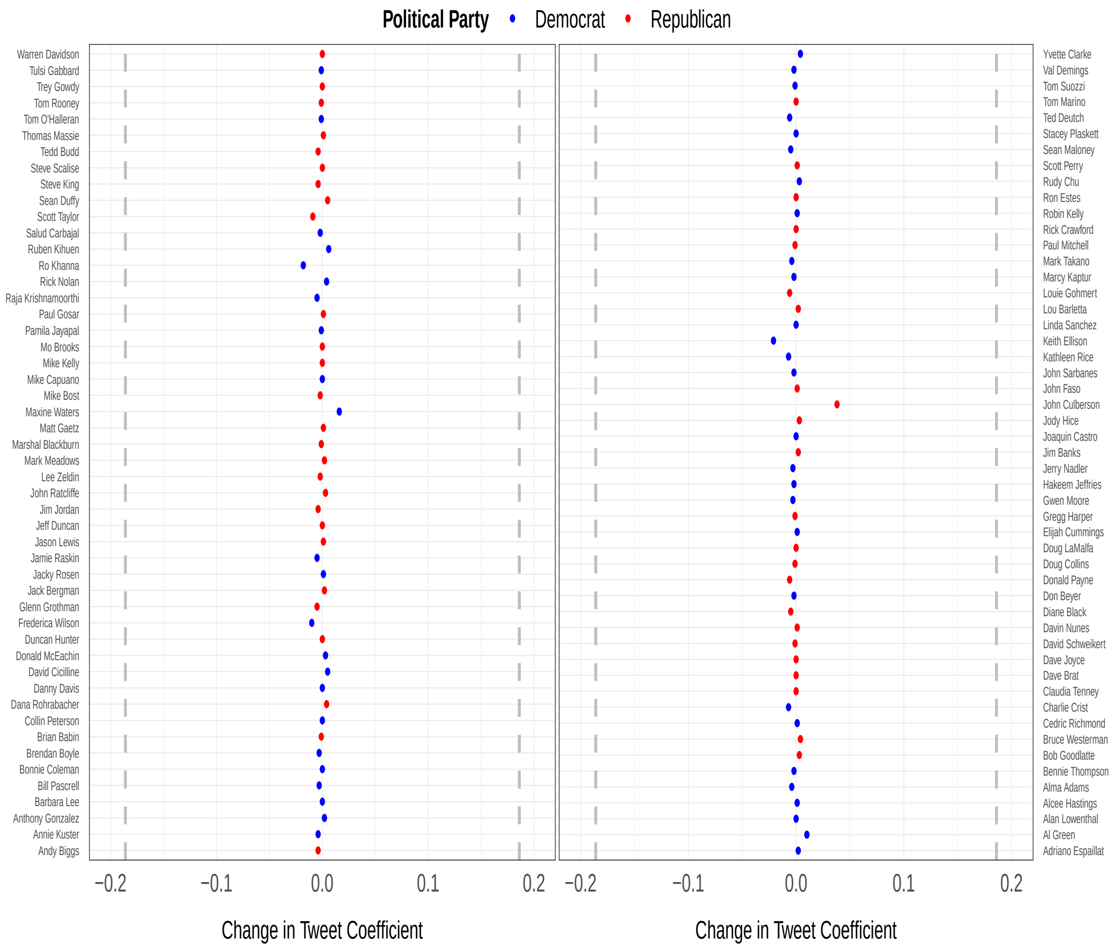
The next, natural, question to ask is whether any of the results are being driven by a small group of legislators? Are a handful of bad apples spoiling the batch? This question is partially addressed by the results presented in the body of the paper. The number of legislators' timelines included in the models falls as the percentages fall. Still, we felt it was necessary to check whether the results are being unduly influenced by any one legislator or by a small group of legislators.

The design of this sensitivity analysis is not as straightforward as one might assume. One approach to the analysis would be to use the timelines of each of the individual legislators as an independent variable. If the individual legislator is driving the results, that legislator's tweets should be significantly related to the hate crimes. The problem with this approach is that our theory suggests that these kinds of models should produce null effects. If we are right, the Twitter activity of all legislators followed by white nationalists can stimulate political violence. If we look at one legislator at a time, we are setting a substantial amount of meaningful variation to zero. The omission of the tweets from the other 106 legislators (Top 25%) in the analysis of a single legislator would omit a majority of the variation and produce a null result.

We approach this sensitivity analysis from the other direction. Rather than omitting 106 members' timelines to analyze the effect of one, we omit the timelines of each legislator individually to see how the omission of each timeline alters the effects. We estimated 107 models, one model for each legislator. If the activity of a single legislator is driving the result, there should be a significant change in the size of the the coefficient for the logged timelines when the activity of that member is removed from the series.

The results for the individual member sensitivity analysis are presented in Figure A.1. The blue and red dots in Figure A.1 represent the change in the coefficients from the coefficient presented in the first model of Table 2 of the main text ($(\hat{\beta}_i - \hat{\beta}_{Main})$). The change associated with the removal of each member is plotted next to their name. The red dots are the differences for the Republican members. The blue dots are the differences for the Democratic members. The gray vertical lines in the plots are the 90% confidence interval for the coefficient from the main model. If any individual

Figure A.1: Change in PEWMA Coefficients by Legislator



Legislator series were constructed by aggregating all the timelines of the top 25 % most followed by white nationalist *excluding* each legislator. The values are the change in the coefficient omitting each legislator (i) from the main model ($\hat{\beta}_i - \hat{\beta}_{Main}$) presented in Table 2. The vertical gray lines are the 90 % confidence interval around the main model coefficient.

legislator is driving the results, the difference between the coefficient from the model that omits the activity of that legislator should be significantly different from the the coefficient from the model that includes that member's activity.

The results presented in Figure A.1 demonstrate that our main results are not being driven by any individual member, or any small group of members, of the house of representatives. None of the differences are statistically significant. In fact, none of the differences even approach standard levels of statistical significance. Most of the differences are grouped around zero. Some are above, some are below, but none of the signs should be interpreted as meaningful because none of the differences can be distinguished from zero. The patterns of the differences also provide further evidence that neither of the parties is driving the results. There are no discernible patterns in the differences in terms of the parties and positive or negative effects. If Republicans or Democrats were driving the results, there would be a propensity for the differences of the different parties to be positive or negative. There is no evidence of this kind of pattern. We conclude that these results validate our choice to combine the timelines of the individual legislators from the individual parties. For the analyses presented in the body of the text, the features of their political profiles that make white nationalists want to follow them are more important than their political parties.

3 Explicit Tweets Hate and Hate Crimes

The final sensitivity analysis is conducted to validate our decision to use the entirety of each member's timeline. We argue that white nationalist perceptions of a member, and the status conferred to the member through their station as a member of the House of Representatives, creates a toxic combination. If the white nationalists that follow you view you as *one of them*, all of your formal actions validate their behavior and beliefs. This is why we use all the tweets from the members' timelines.

There is an alternative causal argument that links Twitter activity of the members in our groups (25%, 10%, and 5%) to hate crimes. It is possible that the only Twitter activity that matters is that is pertinent to white nationalist perspectives. An example can highlight the contrast. Imagine that Steve King sends two tweets. One tweet is about the need of white Americans to have more babies to protect American culture. The other tweet is about abortion. Our argument suggests that both tweets matter to Steve King's white nationalist audience because both tweets are forms of political engagement by someone his white nationalist followers feel affinity for. The alternative argument suggests that only the first tweet is important because it is the only tweet that is relevant to white nationalism. If the latter argument is correct, our approach of aggregating the full timelines could be introducing a lot of irrelevant noise into the variables and underestimating the effects.

Table A.2: PEWMA Models of Twitter Activity, Explicit Tweets, and Hate Crimes

	All Hate Crimes			*Violent Hate Crimes*		
	Top 10%	Explicit	Both	Top 10%	Explicit	Both
Log Tweets	0.313***		0.311***	0.282**		0.282**
	(0.114)		(0.114)	(0.136)		(0.136)
Explicit Tweets		-0.095	-0.094		-0.100	-0.100
		(0.084)	(0.083)		(0.104)	(0.104)
Trump Tweets	0.015+	0.020**	0.016*	0.009	0.013	0.010
	(0.009)	(0.009)	(0.010)	(0.012)	(0.012)	(0.012)
Approval	-0.066	-0.057	-0.066	-0.057	-0.050	-0.057
	(0.068)	(0.069)	(0.068)	(0.068)	(0.069)	(0.068)
Dow Jones	8.609	8.757	7.488	9.977	10.192	8.857
	(9.114)	(9.361)	(9.224)	(11.297)	(11.578)	(11.418)
Charlottesville	-1.245**	-1.194*	-1.268**	-0.808	-0.777	-0.835+
	(0.617)	(0.617)	(0.617)	(0.637)	(0.637)	(0.637)
Charlottesville Week	0.704***	0.735***	0.728***	0.669**	0.698**	0.693**
	(0.256)	(0.256)	(0.256)	(0.298)	(0.298)	(0.298)
Weekends	0.347**	-0.017**	0.337**	0.368**	0.042**	0.360*
	(0.156)	(0.089)	(0.156)	(0.187)	(0.109)	(0.187)
Temperature	0.123**	0.131***	0.126**	0.165***	0.173***	0.167***
	(0.050)	(0.051)	(0.050)	(0.062)	(0.062)	(0.062)
Precipitation	-0.023	-0.020	-0.022	-0.001	0.002	0.000
	(0.038)	(0.038)	(0.038)	(0.046)	(0.046)	(0.046)
$\hat{\omega}$	0.896***	0.893***	0.896***	0.931***	0.93***	0.932***
	(0.022)	(0.023)	(0.023)	(0.019)	(0.02)	(0.02)
AIC	1296.963	1303.206	1297.605	1086.383	1089.752	1087.400
$\text{Wald}_{\hat{\omega}}$	21.379***	22.021***	21.219***	12.487***	12.719***	12.062***
N	411	411	411	411	411	411

Note: *** $p < .01$, ** $p < .05$, * $p < .1$ (two tail); + $p < .1$ (one tail). The columns show the coefficients for the Poisson exponentially weighted moving average (PEWMA) regressions of hate crimes and violent hate crimes. The standard errors are shown in parentheses. The regressions use the logged Twitter timelines from the top 25%, 10%, and 5% of US representatives in terms of identifiable white nationalist followers as a proportion of total followers. The $\hat{\omega}$ coefficient is the estimated discount rate for past observations. The $\text{Wald}_{\hat{\omega}}$ test is a test of the null that $\hat{\omega} = 1$. If $\hat{\omega} = 1$, the PEWMA model reduces to a standard Poisson model.

We tested this alternative argument by coding the timelines of the house members in the top 10% most followed by white nationalists. Coders flagged a tweet as white nationalist if it explicitly made an appeal to White identity or used language to appeal to white nationalists. Steve King's March 12, 2017 tweet, stating the, "[far-right Dutch politician Geert] Wilders understands that culture and demographics are our destiny. We can't restore our civilization with somebody else's babies" is an example of an explicit white nationalist Tweet. This Tweet was *replied* to more than 15,000 times and received more than 11,000 *likes* by other Twitter accounts. An example of language appealing to white nationalist followers is Representative Andy' Bigg's June 29, 2017 tweet, "Kate Steinle. Sarah Root. Grant Ronnebeck. We will no longer be bystanders to these crimes committed by illegal aliens. #SaveAmericanLives." Referring to immigrants as "illegals" dehumanizes them. This is a strategy that should appeal to white nationalists that view immigrants and minorities as less human than white Americans. The tweet also highlights crimes that are part of a popular anti-immigrant narrative. We aggregated these explicit tweets into a separate time series and estimated a separate set of PEWMA regressions like those presented in the main text, one set for all hate crimes and another for violent hate crimes.

The results from our auxiliary analysis using explicit tweets are presented in Table A.2. We present results from six models. The dependent variable in the first three models is the series that includes all hate crimes. The first model shows the results for the top 10% from Table 2 from the main text, the second model shows the PEWMA regression of hate crimes on the explicit tweets, and the third model includes both variables. The dependent variable in the second three models is the series that only includes violent hate crimes. The fourth (10%), fifth (explicit), and sixth (both) models are ordered in the same manner as the first three. The first column lists the independent variables and the quantities of interest.

The results presented in Table A.2 are consistent with the results presented in the body of the manuscript. The explicit tweets variables are not significant in any of the models. The coefficients for the full timelines are not reliably different in the models that include the explicit tweet variables. These results do not provide support for the alternative causal argument. We conclude that our approach is more appropriate than the alternative approach that only uses explicit tweets. The results also support our theoretical argument.

References

Costa, Robert and Ed O'Keefe. 2014. "House Majority Whip Scalise Confirms He Spoke to White Supremacists in 2002." *The Washington Post* .

URL: *https://www.washingtonpost.com/politics/house-majority-whip-scalise-confirms-he-spoke-to-white-nationalists-in-2002/2014/12/29/7f80dc14-8fa3-11e4-a900-9960214d4cd7*$_s$*tory.html?utm*$_t$*erm* $= .514bcdf4a06f$

Gabriel, Trip. 2019. "A Timeline of Steve King's Racist Remarks and Divisive Actions." *The New York Times* .

URL: *https://www.nytimes.com/2019/01/15/us/politics/steve-king-offensive-quotes.html*

Ingraham, Christopher. 2016. "The 'Smoking Gun' proving North Carolina Republicans Tried to Disenfranchise Black Voters." *The Washington Post* July 29.

URL: *https://www.washingtonpost.com/news/wonk/wp/2016/07/29/the-smoking-gun-proving-north-carolina-republicans-tried-to-disenfranchise-black-voters/?utm*$_t$*erm* $= .bfed2b744838$