Politician hate speech and domestic terrorism

James A. Piazza

To cite this article: James A. Piazza (2020) Politician hate speech and domestic terrorism, International Interactions, 46:3, 431-453, DOI: 10.1080/03050629.2020.1739033

To link to this article: https://doi.org/10.1080/03050629.2020.1739033

View supplementary material

Published online: 29 Mar 2020.

Submit your article to this journal

Article views: 121

View related articles

View Crossmark data
ARTICLE

Politician hate speech and domestic terrorism

James A. Piazza

Department of Political Science, The Pennsylvania State University, University Park, PA, USA

ABSTRACT

Does hate speech – rhetoric that targets, vilifies or is intended to intimidate minorities and other groups in society – fuel domestic terrorism? This question is, unfortunately, relevant given the convergence of the use of hate speech by political figures and domestic terrorist incidents in a variety of countries, including the United States. In this study I theorize that hate speech by politicians deepens political polarization and that this, in turn, produces conditions under which domestic terrorism increases. I test this proposition using terrorism and hate speech data for 135 to 163 countries for the period 2000 to 2017. I produce two findings. First, hate speech by political figures boosts domestic terrorism. Second, the impact of political hate speech on domestic terrorism is mediated through increased political polarization.

KEYWORDS

Terrorism; Hate Speech; Political Polarization; Empirical Analysis; Hate Speech; polarization

Does hate speech – rhetoric that targets, vilifies, excludes or is fashioned to intimidate racial, ethnic, religious or sexual minorities, women, political opponents, migrants, disabled persons or members of other groups – by political figures fuel domestic terrorism?¹ This question, unfortunately, gained a new salience in the United States in the wake of the August 3, 2019 mass shooting in El Paso, Texas; an incident that claimed 22 lives and that was investigated by the U.S. Federal Bureau of Investigations as a domestic terrorism incident (Romero and Bogel-Burroughs 2019).

The suspect in the attack, Patrick Wood Crucius, posted an online manifesto minutes before the incident that echoed incendiary public statements made by US President Donald Trump and by other prominent conservative politicians and commentators. For example, in his manifesto Crucius stated that he committed the attack to stem an “Hispanic invasion” of the United States by migrants from Latin American countries. The manifesto vilified Latino immigrants as drains on the U.S. economy, decried the demographic “replacement” of whites by Latinos and other nonwhite minorities, and used the phrase “send them back” with regards to Latinos living in Texas (Arango, Bogel-Burroughs, and Benner 2019).

CONTACT James A. Piazza jap45@psu.edu

¹Terrorism is defined in this study as the deliberate, premeditated use of violence, or the threat of violence, by nonstate actors that is politically motivated and is intended to influence a wider audience beyond individuals affected directly by the attack.

Supplemental data for this article can be accessed on the publisher’s website.

© 2020 Taylor & Francis Group, LLC
Though Crucius denied being motivated by the president – he stated that his anti-Latino and anti-immigrant beliefs predate Trump – journalists, commentators and other politicians were quick to note striking parallels between the shooter’s expressed sentiments and public statements made by Trump and other political figures. An analysis conducted by the *New York Times* after the attack extensively documented the correspondence between Crucius’ expressed motivations in his manifesto and statements about Latinos, nonwhites and immigration made by the president, by Republican politicians and by conservative media commentators. These include almost three dozen instances since 2016 where Trump and others described migration across the U.S.-Mexico border as “an invasion,” Trump’s use of the statement “send them back” in reference to four nonwhite members of Congress who criticized him, the frequent use by Republicans during the 2018 elections of the “immigrant caravan” trope and statements by conservative media figures like Rush Limbaugh and Tucker Carlson predicting that Latino migration will hurt the U.S. labor market, drive down wages, dilute or sully the cultural identity of the U.S., and demographically replace the country’s white population (Peters et al. 2019). Commentators and public officials were quick to condemn Trump’s inflammatory statements about Latinos, Muslims, migrants and minority groups as a dangerous precipitant for violent right-wing extremism. A host of Democratic politicians – former Texas Representative Beto O’Rourke, Senator Bernie Sanders, South Bend, Indiana Mayor Pete Buttigieg, Senator Kamala Harris, Senator Elizabeth Warren, former Vice President Joe Biden – bluntly linked Trump’s rhetoric to the attack and assailed the President for using language that emboldens and encourages domestic terrorists (Villa and Elliott 2019).

The possible connection between use of hateful and discriminatory rhetoric by politicians and domestic terrorist attacks is by no means confined to the El Paso shooting incident. Prior to the El Paso attack, commentators noted similar overlaps between Trump’s rhetoric and the alleged motivations of perpetrators of political violence. For example, in 2016 Patrick Eugene Stein was convicted for plotting to bomb an apartment building lived in and mosque frequented by Somali Muslim refugees in Garden City, Kansas. At his sentencing, Stein’s lawyer pleaded for leniency stating that his client had been influenced by Trump’s frequent statements disparaging Muslims (Farzan 2018). Nor is the convergence of politician hate speech and domestic terrorism limited to the United States. In a variety of countries, discriminatory rhetoric by politicians has been linked to increased domestic terrorist attacks, particularly against local minorities. For example, anti-Muslim rhetoric by politicians during the 2019 parliamentary election campaign in India has been linked to an increase in terrorist attacks by Hindu nationalists against Muslims (Marlow and Chaudhary 2019). Similarly, a bloody 2010 attack on an Ahmadi mosque in Pakistan by the
extremist Salafi Muslim Tehrik-i Taliban terrorist organization has been linked to long-standing hateful rhetoric against Ahmadis by Sunni Islamist politicians (Sayeed 2017). Additionally, the President of Burundi, Pierre Nkurunziza, and members of his ruling political party have been criticized by international organizations and human rights NGOs for use of inflammatory, dehumanizing language in public statements that has been argued to fuel domestic terrorism and political violence (International Crisis Group 2016). Since 2000, politician hate speech, mostly targeting ethnic, racial, social, or religious minorities but sometimes targeting members of rival political groups, has been a feature of domestic terrorism-afflicted countries such as Iraq, Nepal, Somalia, Bangladesh, Turkey, Colombia, Israel, Egypt, Ukraine, Russia, the Philippines, Italy, Greece, Lebanon and Sri Lanka.\(^2\)

In this study I empirically examine the relationship between the use of hateful rhetoric by political figures and domestic terrorism. I produce two significant findings. First, countries featuring more frequent use of hate speech by politicians are more likely to experience an increase in subsequent domestic terrorist attacks. Moreover, the relationship between hate speech and terrorism seems to not be endogenous, rather than endogenous. Second, I find that the impact of hate speech by political parties on domestic terrorism is mediated through increased political polarization. That is, I find that increased hate speech worsens political polarization – defined in the context of the paper as the presence of extreme differences of opinions on political matters among people within the same society\(^3\) – and that this produces a condition under which domestic terrorism is more frequent. As I demonstrate in the next section, there is a scholarly literature linking political polarization to various forms of political violence, including terrorism. However, while this literature shows polarization to be a significant precipitant of political violence, no research has investigated which factors that drive and shape polarization might be important for increasing terrorist activity. The empirical tests I conduct in the study, particularly the mediation analysis, show that political polarization fomented by hate speech by political figures is a crucial and substantial contributor to the occurrence of domestic terrorism. Moreover, by examining the impact of polarization driven by hate speech uttered by politicians or by major political parties in countries, the study underscores the importance of opinion leadership for producing conditions under which deepened polarization increases terrorism.

\(^2\)Source: frequency distributions calculated by the author using data from Mechkova et al. (2019) and the Global Terrorism Database.

\(^3\)Fiorina and Abrams (2008) enumerate significant nuances in the conceptualization and measurement of political polarization. In this study, my conception of political polarization aligns with what Iyengar, Sood, and Lelkes (2012) term “affective polarization” or what (Mason 2015) refers to as “social polarization.” That is, it is both a mass and elite phenomenon involving deep divisions of opinion about a range of political and social topics. This conception of political polarization conforms to the operationalization of the indicator for political polarization used in the analysis (see Mechkova et al. 2019, 21).
Literature Review

Though it is commonplace for commentators to depict political hate speech as a driver of terrorism, as documented above, no systematic cross-national empirical studies have been conducted on the relationship between hate speech and terrorism. However, scholars have examined the impact of hateful, discriminatory and inflammatory rhetoric by politicians on other forms of political violence that share some similarities with terrorism, such as civil wars or hate crimes. Moreover, other strains of empirical research have demonstrated a link between hate speech or violent rhetoric by politicians and increased individual expressions of racism, incivility and aggression, all of which may be important motivators for individuals to engage in political violence. I briefly summarize each of these bodies of work in turn.

Hateful rhetoric by political figures – in particular, rhetoric that targets rival ethnic groups – has been identified as a key precipitant of the onset and intensity of several notable civil wars and genocides. Anti-Croat and anti-Bosnian Muslim rhetoric by Serb politicians in the late 1980s and early 1990s is frequently argued to have been a motivating factor for civil conflict and mass killings by government troops and irregular forces in Former Yugoslavia (Thompson 1999). Dehumanizing depictions of black Africans in Darfur by Sudanese politicians have been correlated with more intense manifestations of civil war violence in Sudan in the mid-2000s (Hagan and Rymond-Richmond 2008). Luoch (2016) argues that hate speech and inflammatory language against rival political-ethnic groups have been an important contributor to political violence in the context of elections during the last 20 years in Kenya. Perhaps the most infamous case of hate speech by political figures fueling mass violence can be found in the Rwandan Civil War in the early to mid-1990s. Multiple scholars allege that anti-Tutsi radio broadcasts by Hutu extremist political figures played a key role in motivating and mobilizing political violence and genocide (see Fujii 2004; Thompson 2007; for a dissenting view see Straus 2007).

A handful of studies indicate a relationship between hate speech and the occurrence of hate crimes. For example, in a qualitative case study of Britain and Germany, Karapin (1999) argued that public statements by politicians that can be interpreted as endorsing or tolerating racist violence prompt perpetration of actual hate crimes against minorities. A recent study by Müller and Schwarz (2018) finds that the prevalence of extremist anti-refugee statements on Facebook by the right-wing Alternative für Deutschland (AfD) political party is associated with increased violent crimes against refugees in areas where social media usage is higher. Chyzh, Nieman, and Webb (2019) examined hateful Twitter speech by U.S. politicians and found it to predict increases in a range of anti-minority incidents, both violent and nonviolent. Multiple studies have

---

4Terrorism differs from these other forms of political violence because it is low-intensity, unconventional and intended to influence an audience rather than to score a battlefield victory – unlike violence that typically is used by insurgent groups in civil wars – and is premeditated rather than impulsive or spontaneous – unlike hate crime.
established a link between President Trump’s rhetoric and hate crimes in the United States. Feinberg, Branton, and Martinez-Ebers (2019) as well as Edwards and Rushin (2018) find that U.S counties that hosted a Trump rally – events at which the president has frequently used incendiary rhetoric against racial and other minorities and political opponents – or that voted for Trump in the 2016 election saw significantly larger increases in hate crimes. Finally, in a study focused on Trump, Müller and Schwarz (2019) found that the President’s tweets on Muslims and on Islam were highly correlated with subsequent anti-Muslim hate crimes.

Moreover, a relatively large number of studies establish a link between hateful, incendiary or violent speech by politicians and increases in racist attitudes, aggression or support for discrimination among individuals. Multiple studies demonstrate that subjects who are exposed to racially charged or prejudiced statements by President Trump are more likely themselves to subsequently express racist sentiments against minorities and others (Crandall, Miller, and White 2018; Schaffner 2019). Kalmoe (2014) and Kalmoe, Gubler, and Wood (2018) found that use of violent metaphors by politicians increases support for political violence among aggressive individuals and among strong partisans. Survey experiments conducted among Israeli and Indian subjects determined that politician use of violent rhetoric, specifically mild fighting words, in the context of disparaging racial or religious outgroups increased support for public policies that are deleterious for the outgroups targeted by the rhetoric (Gubler and Kalmoe 2015).

Though these studies do provide preliminary evidence that politician hate speech stokes forms of political violence, they differ from my study in important ways. First, as previously noted, they examine forms of political violence other than terrorism. Second, previous research on hate speech and political violence is limited to individual cases where hateful rhetoric by politicians spurred violent civil conflicts, hate crimes, contentious politics or discriminatory policies that targeted vilified outgroups. As such, they are not able to establish a generalizable pattern whereby hate speech by political figures promotes political violence in a variety of contexts. Finally, and perhaps most importantly, while some of the studies do speculate on how hate speech might prompt political violence, none of them identify and test a specific process through which hateful rhetoric by politicians stokes political violence. In the next section I argue that political polarization is an important mechanism that mediates the relationship between political hate speech and terrorism.

**Hate Speech, Polarization and Terrorism**

If hate speech is indeed a precipitant of domestic terrorism, how might hateful rhetoric by politicians actually foment domestic terrorist activity within
countries? I argue that when politicians employ hateful, bigoted and discriminatory rhetoric against groups within society – typically as a means to mobilize their political supporters and to neutralize their opponents and deflect critics (see Zeitzoff 2018) – an important byproduct of hate speech is produced: increased political polarization. When politicians employ hate speech to demonize political, social, ethnic, cultural or religious groups in society, they prompt members of those groups to increase their affiliation with individuals who are of the same political, social, ethnic, cultural or religious background – ingroup members – and to decrease their tolerance and acceptance of individuals who are members of other rival or disparaged political, social, ethnic, cultural or religious outgroups. Much in the same way that perception of threat has been theorized to increase negative attitudes toward outgroups and affinity within ingroups (see, for example, Brewer 1999), I argue that hate speech and the strong emotions it arouses fosters “ingroup love and outgroup hate.” In essence, hate speech helps to deepen and fuel political tribalism. This has two effects related to terrorism. First, hate speech facilitates the dehumanization of outgroup members which makes it easier for militants to commit acts of political violence against them. Increased outgroup hate reduces inhibitions of perpetrators of political violence while mitigating public backlash, at least from other ingroup members. This produces a condition under which terrorism is a less politically or socially costly behavior. Second, by encouraging prejudicial and hostile attitudes against outgroups, hate speech, particularly by powerful political figures and parties, can foster greater unity and conformity of opinion among ingroup members. Hateful rhetoric often convinces its intended target population that they are superior to hated outgroups, that rival ethnic, social, political, cultural or religious communities are hostile and threatening and that now is the time to reinforce their own community bonds to counter rivals. By producing this condition, hate speech legitimizes and builds community support for affiliated extremist political actors, producing resources, recruits, sympathizers and apologists for militant activities. The end result is increased terrorism.

Several literatures are broadly consistent with this theoretical story. The argument that hate speech and other types of extreme or violent rhetoric by political figures increases polarization is well documented by scholars (Gubler and Kalmoe 2015; Kalmoe, Gubler, and Wood 2018; Kalmoe, Gubler, and Wood 2014; Müller and Schwartz 2018; Sunstein 2002; Zeitzoff 2018). It is also consistent with parallel bodies of research that find hate speech and extremist rhetoric to: reinforce ingroup identification while increasing prejudices against outgroups (Soral, Bilewicz, and Winiewski 2018); create a permissive and normalized environment for discriminatory attitudes against outgroups (Chyzh Nieman, and Webb 2019; Crandall Miller, and White 2018; Javed and Miller 2019); facilitate the formation of “harmful social movements” that engage in violent collective action (Tsesis 2002); and to reduce inhibitions against using violence against hated outgroups (Bandura et al. 1996; Dharmapala and McAdams 2005).
Polarized societies are, in turn, more likely to experience political violence and terrorism, of all types. This is borne out by a number of studies. Historical investigations of the collapse of the Weimar Republic in Germany credit extreme political polarization with the outbreak of political violence and acts of terrorism (Bracher 1955; Winkler 1994). Linz (1978) makes similar observations for various Latin American regimes in the 1960s and 1970s. Social and political polarization has been also found to be important for fueling civil wars in countries (Elbadawi 1999; Gubler and Selway 2012; Humphreys and Weinstein 2008; Montalvo and Reynal-Querol 2005; Scarcelli 2014). Similarly, Balcı (2010) found that during the Spanish Civil War pre-conflict political polarization determined which parts of the country experienced more intense violence. Experimental research on individual subjects by Kalmoe and Mason (2019) provides a useful explanation for why political polarization might drive political violence. The authors find that highly partisan individuals in politically polarized societies are more likely to tolerate, rationalize and in cases even endorse violence against political opponents.

Polarization has also been linked specifically with terrorist activity. Python, Brandsch and Tshkay (Python, Brandsch, and Tshkay 2017) found terrorism to be more frequent in countries and during time periods characterized by higher levels of political polarization. An empirical study conducted by O’Brien and Haider-Markel (1998) found political polarization to be associated, albeit weakly, with right wing militia activity in the United States. Periods of increased political polarization in Europe and Latin America in the 1960s were found by Martin and Perliger (2012) to coincide with increased terrorism. The authors theorized that in these cases political polarization rendered the use of violence by militant groups more acceptable to a wider swath of the public. Sayari (2010) produced a similar finding for the incidence of left-wing terrorism in Turkey in the 1970s and 1980s. Finally, Weinberg, Pedahzur, and Perliger (2008) observed that political polarization, in part, prompted political parties in several countries to turn to terrorism.

**Hypotheses**

Given my theoretical assumptions and the corresponding literature, I have two expectations. First, I expect countries where hate speech by politicians is more pervasive to experience higher levels of domestic terrorism.

---

Note, in the studies cited here, political polarization is linked to increased domestic terrorism of all ideological types. Political polarization is associated with societal ideological clashes that prompt individuals and groups to engage in terrorism against political opponents, outgroups and other demonized targets. As a consequence, my expectation is that political polarization boosts domestic terrorist incidents generally. To empirically check this, I ran naive estimations regressing the measure of political polarization within countries to counts of right-wing, left-wing, nationalist-separatist and religious-Islamist terrorism. I found that political polarization is a significant, positive predictor of all of these types. (Results available from the author.) This gives me further assurance that political polarization is associated with increases of terrorism in general, and of all types.
This is expressed in hypothesis 1: $H_1$ Countries where political figures use hate speech more frequently experience more domestic terrorist attacks. Second, I argue that it is by increasing political polarization within countries that hate speech by politicians drives higher rates of terrorism. This expectation is expressed in hypothesis 2: $H_2$ The impact of hate speech by political figures on terrorism is mediated by increased political polarization.

In the next section, I present the research design I use to test both of these hypotheses.

**Research Design**

To investigate the impact of hate speech by politicians on domestic terrorism, and to assess the theorized channel – increased polarization – through which hate speech might prompt domestic attacks, I conduct two sets of empirical tests on a cross-sectional, time series sample of between 135 and 163 countries for the period 2000 to 2017.\(^6\) The first set of analyses employs negative binomial count estimations, along with a two-stage least squares estimation to test for endogeneity, while the second uses a test of mediation to determine whether the effect of hate crime on domestic terrorism is mediated by increased political polarization.

The dependent variable for the analyses is measured as a count of domestic\(^7\) terrorist attacks occurring within a country. The source for the dependent variable is the Global Terrorism Database (GTD) which is produced and curated by the National Consortium for the Study of Terrorism and Responses to Terrorism at the University of Maryland.\(^8\) Because terrorist attacks in countries are counts that cannot produce negative values, are over-dispersed and are likely not independent of one another in terms of observations, I use a negative binomial rather than ordinary least squares technique when conducting the empirical analysis (see Hilbe 2011).\(^9\)

In theory, hate speech by politicians should stimulate domestic terrorism of all types against a range of targets or victims. Political hate speech might prompt partisans to launch terrorist attacks against victims associated with their political rivals. Hate speech could mobilize members of an ethnic, racial, religious or cultural majority to commit terrorist attacks against various minority group targets or against despised outgroup members. Hate speech

---

\(^6\)The temporal limitations of the study are set by data availability. Data on terrorist attacks, the dependent variable, are available only through 2017. Data on hate speech, the independent variable, is available only from 2000 on.

\(^7\)A domestic terrorist attack is defined as a terrorist attack occurring with the boundaries of one country that is perpetrated by a citizen or resident of the country against domestic targets or victims.

\(^8\)Data and codebook for the GTD can be found at: [https://www.start.umd.edu/gtd/](https://www.start.umd.edu/gtd/).

\(^9\)To check for outlier effects, I collapsed the counts of terrorist attacks into a dichotomous measure coded 1 for observations containing one or more attack and zero for observations containing no terrorist attacks. I then reran the main model (model 2) using a logistical regression technique. This produced the same substantive results as the main model. Results summarized in Appendix Table 2.
could also create a climate of alienation, discrimination and oppression that foments strong grievances among minorities and outgroups, prompting members of those communities to engage in terrorism. Therefore, I expect hate speech to increase domestic terrorism generally, among a range of perpetrators and against a range of targets or victims, in countries where it occurs.

The independent variable for both analyses, use of hate speech by politicians, is derived from the “Political Parties Hate Speech” indicator\textsuperscript{10} developed by the Digital Society Project (Mechkova et al. 2019) and published in the Varieties of Democracy (V-Dem) database. This variable is measured on an ordinal scale and captures how often political parties within a country employ hate speech as part of their public statements, speeches and propaganda. The hate speech variable is limited to public statements and official rhetoric used by major political parties within countries. It excludes, therefore, fringe political parties or political movements, pressure groups and individuals that are not within the mainstream arena of political contestation. Hate speech is defined by (Mechkova et al. 2019, 26) as, “... any speech that is intended to insult, offend, or intimidate members of specific groups, defined by race, religion, sexual orientation, national origin, disability or similar trait.” The original hate speech measure is coded between zero, indicating that major political parties employ hate speech “extremely often,” to 4, indicating parties use hate speech “never, or almost never.”\textsuperscript{11} To simplify interpretation, I inverted this scale so that higher scores indicate more frequent use of hate speech by political parties.\textsuperscript{12} Use of hate speech by

\textsuperscript{10}Variable: "v2smpolhate_ord" from (Mechkova et al. 2019, 26). The question used to develop this variable reads, “How often do major political parties use hate speech as part of their rhetoric?” Note, it would be useful to compare the Political Parties Hate Speech indicator – a measure developed by country expert ratings, as explained in the next footnote – with some sort of count or content measure of actual hate speech messages from politicians to improve confidence in the findings. However, to my knowledge, no cross-national database of hate speech messages yet exists. This would be a very useful tool for future study of the relationship between hate speech and terrorism.

\textsuperscript{11}The hate speech measure developed by Mechkova et al. (2019) is produced using the V-Dem expert panel coding system. Coppedge, et al. (2019) explain this process in detail. Evaluative indicators in V-Dem are produced using multiple ratings by groups of country experts. Country experts are recruited by V-Dem staff for each country in the database through a rigorous process that involves competitive selection from a larger pool based upon the experts’ backgrounds, levels of expertise, diversity of experience, and assessment of their impartiality. Around 60 percent of country experts are nationals or permanent residents of the country they evaluate. For each indicator, a minimum of five country experts provide an independent rating. Country experts provide a rating for each country-year observation independently. For ordinal-level indicators like those used in this study, experts are provided detailed vignettes (similar to rubrics) describing each ordinal level to allow the experts to match their rating decision to the situation in the country for that year. Moreover, to ensure that expert ratings are consistent across countries and years, V-Dem employs “bridge coding” whereby some experts produce ratings across multiple countries, or multiple years, for an indicator. These bridge codes are then used to check the ratings of experts who have coded an indicator for only one country or one year at a time. Once experts submit their ratings, V-Dem combines them using a an IRT (item response theory) model that takes into account patterns of cross-expert agreement and disagreement. Specifically, the expert ratings are combined into an, “... integerized median ordinal highest posterior probability category over measurement model output (Coppedge et al. 2019, 143).” Also, V-Dem has experts fill out a post-survey questionnaire which is used to identify sources of attitudinal or other biases.

\textsuperscript{12}Major political parties use hate speech: 4 = extremely often; 3 = often; 2 = sometimes; 1 = rarely; 0 never or almost never.
political parties is not infrequent in the sample. In around 15 percent of observations politicians are rated as using hate speech “often” or “extremely often.” Use of hate speech occurs “sometimes” in about 33.4 percent of observations. Hate speech is “rare” in 34.8 percent of observations and is reported as “never” or “almost never” occurring in about 16.4 percent of observations.

For the mediation test, the details of which are explained below, I utilize another indicator developed by Mechkova et al. (2019) that operationalizes the level of political polarization for each country during each year observed. This indicator\textsuperscript{13} rates polarization using an ordinal scale ranging from zero, indicating “serious polarization” where there are “serious differences in opinions in society on almost all key political issues, which result in a major clash of views” to 4, indicating “No polarization. There are differences in opinion but there is a general agreement on the directions for key political issues (see Mechkova et al. 2019, 21).” Mechkova (2019, 27) and her colleagues clarify that polarization is not mere the existence of a plurality of opinions within a country on political or social issues. Rather, polarization is defined as a situation where differences in opinions lead to “major clashes of views” and sharp disagreement among people on the “general direction” society should follow. As was the case for the hate speech indicator, I inverted the ordinal scale for the polarization measure so that higher scores on the 0 to 4 scale indicate observations of higher levels of polarization.\textsuperscript{14}

In the sample, around 23.8 percent of the country-year observations are characterized as exhibiting severe polarization, around 30.8 percent exhibit moderate polarization, around 23.1 percent medium polarization; around 20.7 exhibit limited polarization and 1.3 percent are coded as having no polarization.

Both hate speech and polarization are broadly distributed across the country-year sample, but there are some basic patterns. Appendix Figures 1, 2 and 3 help to illustrate these patterns. In terms of regime type, democracies are slightly less likely than average to experience both hate speech and political polarization while anocracies exhibit higher levels of both hate speech and polarization than other regime types. Interestingly, dictatorships produce higher levels of hate speech than average – though lower than anocracies – but are characterized by lower levels of polarization. This might be due to the higher degree of control over independent political behavior that dictatorships wield. In terms of world geographic region, hate speech is most acute in South Asia, Sub-Saharan Africa, Eastern Europe and Central Asia and the Middle East and North Africa. It is lower than the world average in regions like Latin America, Southeast Asia, East Asia and Western

\textsuperscript{13}Variable “v2smpolso.ord” from (Mechkova et al. 2019, 21). Question: “How would you characterize the differences of opinion on major political issues in society?”

\textsuperscript{14}4 = serious polarization; 3 = moderate polarization; 2 = medium polarization; 1 = limited polarization; 0 = no polarization.
Europe and North America. Polarization is most acute in Latin America, South Asia, Southeast Asia and Sub-Saharan Africa. It is lower than average in the Middle East and North Africa, Western Europe and North America and East Asia. Finally, levels of hate speech and polarization have modestly increased since 2000. Between 2000 and 2018, the frequency and intensity of hate speech by political parties across all countries increased by 21.7 percent. Polarization increased by 14.1 percent across all countries for the same time period.

To check against spuriousness, I also include some control variables in the models. I control for government efforts to censor the media (Pemstein et al. 2019)\(^{15}\) to hold constant and separate out the environment for all permitted political speech from hate speech in particular. I also control for political regime type by including the Polity 2 indicator and a squared version of Polity 2 (Marshall, Gurr, and Jaggers 2018). Research by Gaibulloev, Piazza, and Sandler (2017) demonstrates that terrorist attacks are significantly more frequent in countries with anocratic or hybrid political regimes and are less likely to occur in established democracies and autocratic regimes. To hold constant the impact of political stability, or change, on terrorism, I include an indicator for the age of the political regime in the country, using the “durable” indicator from Polity (Marshall, Gurr, and Jaggers 2018). Some studies have found that younger regimes, particularly young democracies, are more terrorism prone (for example, Piazza 2013). I also control for standard country level features such as national population and level of economic development. To do this I use natural log transformed measures of population and gross domestic product per capita.\(^{16}\) More populous countries are regularly found to feature higher rates of terrorism while level of economic development has been found to affect patterns of terrorist attacks in countries (see Gassebner and Luechinger 2011). I control for the degree of ethno-linguistic fractionalization in the country.\(^{17}\) Countries that are highly fractionalized along ethnic or linguistic lines might experience more terrorism, or might affect the impact of hate speech on terrorism. I also control for the general level of internal armed political violence within the country, using a count of the number of internal armed conflicts occurring during the observation.\(^{18}\) Of course, countries experiencing large numbers of internal armed conflicts and civil wars are expected to also experience higher levels of terrorism. In all models I also include the number of terrorist attacks

---

15 Derived from “Government Censorship Effort – Media” (v2mecenefm) from Pemstein et al. (2019, 21). To ease interpretation, I invert the ordinal scale for government censorship of media so that 4 = government “attempts to censor are direct and routine,” 3 = attempts are “indirect but nevertheless routine,” 2 = attempts are “direct but limited to especially sensitive issues,” 1 = attempts are “indirect and limited to sensitive issues,” and 0 = “government rarely attempts to censor major media in any way.”


17 Source: Alesina et al. (2003). Data obtained from the Quality of Government Database, which provided updated data for entire timeseries. Variable=“al_ethnic.”

18 Source: UCDP/PRIO. Variable = “ucdp_type3, Internal Armed Conflict.”
that occurred in the previous year in the country. This helps to control for terrorist attack spells and also addresses temporal autocorrelation. Terrorist campaigns frequently cluster temporally within countries and can occur over years.\textsuperscript{19} All independent and control variables are lagged by one year as the impact of hate speech on terrorism might not be immediate.\textsuperscript{20}

Descriptive statistics for all variables used in the study can be found in Table 1.

\textbf{Results}

The results of the analysis demonstrate that hate speech by politicians does, indeed, increase the amount of domestic terrorism that a country experiences. This provides support for the first hypothesis. Table 2 summarizes this finding. In model 1, I conduct a semi-naïve estimation that includes only the hate speech variable and the measure of the previous year’s terrorist attacks. Model 2 is the full estimation, including all covariates. In model 3, I add political polarization as a covariate into the estimation. Political polarization is significant in this model and its inclusion reduces the size of the coefficient for hate speech. This is a preliminary indication that the impact of hate speech on domestic terrorism could be mediated through increased political polarization. I further examine this using mediation tests, discussed below. In model 4, I rerun the full estimation using country-year fixed effects to control for country-level idiosyncracies; models 1 through 3 include standard errors clustered on the country to adjust for autocorrelation within units in the panel. Finally, in model 5 I rerun the full estimation but drop the lagged dependent variable. In all estimations, hate speech is a positive, significant predictor of subsequent domestic terrorist attacks. For each ordinal level increase in the frequency of use of hate speech by politicians and political parties, the number of domestic terrorist attacks increases by between 1.8 attacks (model 2) and 3.34 attacks (model 1).\textsuperscript{21} The substantive impact of hate speech on domestic terrorism is comparable to that of population size and the prevalence of intrastate war. This is represented in Figure 1, which graphs the marginal effects of all variables in the analysis.\textsuperscript{22}Figure 2 presents the substantive overall model effect of the full estimation represented in model 2. It demonstrates that, holding all other covariates (censorship, regime type, regime durability, population, GDP per capita, experience of intrastate conflict and experience of past terrorism) constant at their median or

\textsuperscript{19}Note, as a check I reran all models excluding terrorist attacks from the previous year. These produce the same results as those in the main analysis. Results available from the author.

\textsuperscript{20}As a check, I also ran the estimations without lagging the independent variables and with lagging them by two years instead of one. These reproduced the main findings of the study. I have published the results for the main model (model 2) in Appendix Table 3. (Full results available from author.)

\textsuperscript{21}Calculated using post-estimation marginal effects simulations of a one-unit increase of hate speech while holding all other covariates constant at their mean values. Note, the substantive impact of a one-unit increase in hate speech in model 4, which excludes the lagged dependent variable as a predictor, is 3.071 more domestic terrorist attacks.

\textsuperscript{22}Note, Figure 1 presents the marginal effects simulations for the main estimation: model 2.
mean values, as political parties within countries increasingly use hate speech, the country experiences higher levels of terrorism. For example, countries where political parties “never or almost never” engage in hate speech are projected to experience on average around .9 domestic terrorist attacks per year, holding all other factors constant. Countries characterized by “rare” use of hate speech by parties are projected to experience around 1.6 attacks annually while countries where political parties “sometimes” use hate speech are expected to feature 2.9 incidents of domestic terrorism per year. In countries rated by Mechkova et al. (2019) as having political parties that “often” hate speech, around 5.2 domestic terrorist attacks are expected to occur. Finally, in countries where political parties use hate speech “extremely often,” the model projects around 9.5 domestic terrorist attacks annually. For each ordinal scale increase in the hate speech variable, the projected number of domestic terrorist nearly doubles.

Endogeneity

It is possible that hate speech and domestic terrorism are endogenously related. In additional to precipitating domestic terrorism, hate speech by political figures might also increase in the wake of terrorist attacks within a country. To address this possibility, I conduct two-stage least squares estimations instrumenting for hate speech by political parties with

### Table 1. Descriptive statistics.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs.</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domestic Terrorist Attacks</td>
<td>2,742</td>
<td>28.85157</td>
<td>161.154</td>
<td>0</td>
<td>2809</td>
</tr>
<tr>
<td>Hate Speech by Political Parties</td>
<td>2,742</td>
<td>1.442378</td>
<td>.9228534</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Polarization</td>
<td>2,742</td>
<td>2.551058</td>
<td>1.105251</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Gov. Censorship of Media</td>
<td>2,742</td>
<td>1.505835</td>
<td>1.320482</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Regime Type (Polity 2)</td>
<td>2,742</td>
<td>4.052881</td>
<td>6.201509</td>
<td>−10</td>
<td>10</td>
</tr>
<tr>
<td>Polity 2 squared</td>
<td>2,742</td>
<td>54.87053</td>
<td>33.96116</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>Regime Durability</td>
<td>2,742</td>
<td>27.31911</td>
<td>31.72185</td>
<td>0</td>
<td>208</td>
</tr>
<tr>
<td>(ln) GDP Per Capita</td>
<td>2,742</td>
<td>8.414064</td>
<td>1.535095</td>
<td>5.40514</td>
<td>11.62606</td>
</tr>
<tr>
<td>Interstate Conflict (UCDP)</td>
<td>2,742</td>
<td>.2140773</td>
<td>.7464709</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>Weak anti-Defam/Hate Speech Leg.</td>
<td>2,742</td>
<td>1.619985</td>
<td>1.063134</td>
<td>0</td>
<td>4</td>
</tr>
</tbody>
</table>

It is also possible that hate speech and polarization are endogenous to one another. Statistically, they are correlates ($\rho = .590$). In the paper, I theorize that hate speech increases polarization to produce conditions under which domestic terrorism is more likely. However, it is possible that hate speech by politicians is more common, and acute, in environments that are already highly polarized, and that this might complicate the relationship between hate speech and terrorism. I empirically check for this and do not find it to be the case. I split the sample into two sets of observations characterized by “low” and “high” polarized environments. The former consists of observations where polarization is completely absent or is “limited.” The latter consists of observations where polarization is “medium,” “moderate” or “serious.” I then reran the main estimations on the two sets. I found that hate speech drives domestic terrorism in both “low” and “high” polarized environments. To check further, I split the sample again and examined the impact of hate speech on terrorism in countries with the highest two ordinal categories of polarization: “moderate” and “serious.” I found, again, that in these highly polarized environments, hate speech increased domestic terrorism. These results are summarized in Appendix Table 5. The results give me confidence that though hate speech and polarization are possibly related, the impact of hate speech on terrorism is not due to a wider environment of elevated polarization.
Table 2. Main results.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Y:</td>
<td>Domestic Terrorist Attacks</td>
<td>Domestic Terrorist Attacks</td>
<td>Domestic Terrorist Attacks</td>
<td>Domestic Terrorist Attacks</td>
<td>Domestic Terrorist Attacks</td>
</tr>
<tr>
<td>Hate Speech by Political Parties$_{t-1}$</td>
<td>0.573***</td>
<td>0.598***</td>
<td>0.378**</td>
<td>0.158**</td>
<td>0.479**</td>
</tr>
<tr>
<td></td>
<td>(0.117)</td>
<td>(0.128)</td>
<td>(0.138)</td>
<td>(0.053)</td>
<td>(0.165)</td>
</tr>
<tr>
<td>Political Polarization$_{t-1}$</td>
<td></td>
<td></td>
<td>0.386**</td>
<td>0.255***</td>
<td>0.481***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.118)</td>
<td>(0.047)</td>
<td>(0.145)</td>
</tr>
<tr>
<td>Gov. Censorship of Media$_{t-1}$</td>
<td>−0.128</td>
<td>−0.179</td>
<td>−0.104*</td>
<td>−0.354*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.131)</td>
<td>(0.126)</td>
<td>(0.046)</td>
<td>(0.175)</td>
<td></td>
</tr>
<tr>
<td>Regime Type (Polity 2)$_{t-1}$</td>
<td>−0.002</td>
<td>−0.019</td>
<td>0.008</td>
<td>−0.052</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.029)</td>
<td>(0.029)</td>
<td>(0.011)</td>
<td>(0.040)</td>
<td></td>
</tr>
<tr>
<td>Polity 2 squared$_{t-1}$</td>
<td>−0.010*</td>
<td>−0.008</td>
<td>−0.000</td>
<td>−0.019**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.005)</td>
<td>(0.002)</td>
<td>(0.006)</td>
<td></td>
</tr>
<tr>
<td>Regime Durability$_{t-1}$</td>
<td>−0.003</td>
<td>−0.002</td>
<td>−0.007***</td>
<td>−0.007</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.002)</td>
<td>(0.006)</td>
<td></td>
</tr>
<tr>
<td>(ln) Population$_{t-1}$</td>
<td>0.701***</td>
<td>0.694***</td>
<td>0.263***</td>
<td>0.983***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.093)</td>
<td>(0.092)</td>
<td>(0.034)</td>
<td>(0.116)</td>
<td></td>
</tr>
<tr>
<td>(ln) GDP Per Capita$_{t-1}$</td>
<td>0.328**</td>
<td>0.280*</td>
<td>0.268***</td>
<td>0.500***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.125)</td>
<td>(0.116)</td>
<td>(0.044)</td>
<td>(0.122)</td>
<td></td>
</tr>
<tr>
<td>Ethno-Ling. Fract$_{t-1}$</td>
<td>−0.414</td>
<td>−0.534</td>
<td>−0.030</td>
<td>−0.481</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.457)</td>
<td>(0.455)</td>
<td>(0.193)</td>
<td>(0.627)</td>
<td></td>
</tr>
<tr>
<td>Interstate Conflict (UCDP)$_{t-1}$</td>
<td>0.592***</td>
<td>0.559***</td>
<td>0.139***</td>
<td>0.686***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.134)</td>
<td>(0.133)</td>
<td>(0.026)</td>
<td>(0.181)</td>
<td></td>
</tr>
<tr>
<td>Terrorist Attacks$_{t-1}$</td>
<td>0.018***</td>
<td>0.008*</td>
<td>0.007*</td>
<td>0.001***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.003)</td>
<td>(0.000)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.381*</td>
<td>−13.115***</td>
<td>−13.221***</td>
<td>−8.068***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.223)</td>
<td>(1.877)</td>
<td>(1.857)</td>
<td>(0.759)</td>
<td></td>
</tr>
<tr>
<td>Obs.</td>
<td>2,744</td>
<td>2,550</td>
<td>2,550</td>
<td>2,226</td>
<td></td>
</tr>
<tr>
<td>Wald χ²</td>
<td>63.63***</td>
<td>247.97***</td>
<td>270.00***</td>
<td>458.99***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>271.54***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pseudo r²</td>
<td>0.0828</td>
<td>0.1239</td>
<td>0.1281</td>
<td>0.1038</td>
<td></td>
</tr>
<tr>
<td>Clusters (countries)</td>
<td>163</td>
<td>153</td>
<td>153</td>
<td>153</td>
<td></td>
</tr>
<tr>
<td>Model</td>
<td>NBREG</td>
<td>NBREG</td>
<td>NBREG</td>
<td>XTNBREG</td>
<td></td>
</tr>
<tr>
<td>Country and Year Fixed Effects</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

* p ≤ .1  ** p ≤ .01  *** p ≤ .000.
a variable measuring the absence or weakness of government online defamation and hate speech laws. Governments that lack a legal framework to prosecute online defamation and hate speech create an environment where hate speech by political parties is more likely. However, lack of such legal framework should have no direct impact on terrorism, independent of increasing the prevalence of hate speech itself.\textsuperscript{24} I derive this instrument using an indicator\textsuperscript{25} from Mechkova et al. (Mechkova et al. 2019, 21) that is a zero to four-point ordinal scale measuring the degree to which a country’s, “… legal framework provide[s] protection against defamatory online content, or hate speech … ”\textsuperscript{26} In around 16.9 percent of the observations, countries offer full legal protections against defamation or hate speech. In 27.7 percent of the observations, protection is mostly offered. In 31.7 percent, the legal system provides some protection. In nearly 24 percent of observations, protection against defamatory content or hate speech is either not really or not provided by the legal system at all.

\textsuperscript{24}Statistical checks provide support for this logical assumption. While weak anti-defamation or hate speech laws in countries statistically predict subsequent hate speech, such laws do not statistically predict subsequent domestic terrorist attacks. This is demonstrated in Appendix Table 4.

\textsuperscript{25}Mechkova (2019, 21) "Defamation protection" (v2smlawpr).

\textsuperscript{26}To ease interpretation, I inverted the scale so that 4 = No. The law provides no protection against Internet defamation and hate speech; 3 = Not really; 2 = Somewhat; 1 = Mostly; 0 = Yes.
The results of the endogeneity tests are summarized in Table 3. In both model 5, where the semi-naïve version of the estimation is run, and model 6, where the full estimation is run, the instrumented independent variable is a significant positive predictor of domestic terrorism. The Durbin chi-square and Wu-Hausman F tests are also significant. This suggests that hate speech by political parties is causally associated with an increase in domestic terrorism in countries, and that the instrument used in the endogeneity tests – weak anti-defamation and hate speech legal framework – is an instrument satisfies the exclusion restriction.

Polarization as a mediator

I also find evidence that hate speech by political parties boosts domestic terrorism by stoking more intense political polarization within society. This provides support for the second hypothesis and is illustrated in Figure 3.27 To conduct the mediation test, I used a statistical package developed by Hicks and Tingley (2011).28 Because this package cannot accommodate negative binomial estimations and has been designed for ordinary least squares (OLS) or logistical/probit regression techniques,

Figure 2. Impact of political party use of hate speech on projected incidents of terrorism, 2000–2017. Notes: Sample: 156 countries for the years 2000–2017; 2,591 total observations. Marginal effects simulations after negative binomial regression estimation included as controls: government censorship of media; political regime type and regime type squared; age of political regime; national population; gross domestic product per capita; number of intrastate wars occurring in country; count of previous year’s terrorist attacks.

The results of the endogeneity tests are summarized in Table 3. In both model 5, where the semi-naïve version of the estimation is run, and model 6, where the full estimation is run, the instrumented independent variable is a significant positive predictor of domestic terrorism. The Durbin chi-square and Wu-Hausman F tests are also significant. This suggests that hate speech by political parties is causally associated with an increase in domestic terrorism in countries, and that the instrument used in the endogeneity tests – weak anti-defamation and hate speech legal framework – is an instrument satisfies the exclusion restriction.

Polarization as a mediator

I also find evidence that hate speech by political parties boosts domestic terrorism by stoking more intense political polarization within society. This provides support for the second hypothesis and is illustrated in Figure 3.27 To conduct the mediation test, I used a statistical package developed by Hicks and Tingley (2011).28 Because this package cannot accommodate negative binomial estimations and has been designed for ordinary least squares (OLS) or logistical/probit regression techniques,

27 To produce the results in Table 4, I used a Stata package (“medeff”) developed by Hicks et al. (2011).
28 “medeff” (Hicks and Tingley 2011), which is based upon the technique for mediation developed in Imai, Keele and Tingley (2010) and Imai, Keele, Tingley and Yamamoto (2010).
I transformed the dependent variable into a dichotomous measure coded 1 for any observation with one or more attacks. This is appropriate given the rarity of terrorist attacks in the sample and the fact that the first mode for the dependent variable is zero and the second mode is a value of one attack; together the zero and one values for the terrorism measure account for nearly 70 percent of all observations.29 However, as a check I also ran the mediation test using an OLS estimation on the count measure. This produced the same overall result, that the impact of hate speech on domestic terrorism is significantly mediated through increased political polarization.30 In Figure 3, hate speech is found to be a significant predictor of polarization.31 Polarization, in turn, is a significant predictor of domestic terrorism.32 Moreover, when hate speech is regressed to terrorism without polarization, it is a significant predictor of terrorism.33 However, when polarization is

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Hate Speech by Political Parties$_{t-1}$</td>
<td>6.708* (2.966)</td>
<td>10.929* (6.361)</td>
</tr>
<tr>
<td>Gov. Censorship of Media$_{t-1}$</td>
<td>−4.807* (2.156)</td>
<td></td>
</tr>
<tr>
<td>Regime Type (Polity 2)$_{t-1}$</td>
<td>−0.532 (0.379)</td>
<td></td>
</tr>
<tr>
<td>Polity 2 squared$_{t-1}$</td>
<td>−0.028 (0.051)</td>
<td></td>
</tr>
<tr>
<td>Regime Durability$_{t-1}$</td>
<td>0.003 (0.047)</td>
<td></td>
</tr>
<tr>
<td>(In) Population$_{t-1}$</td>
<td>1.710* (0.768)</td>
<td></td>
</tr>
<tr>
<td>(In) GDP Per Capita$_{t-1}$</td>
<td>1.731 (1.260)</td>
<td></td>
</tr>
<tr>
<td>Interstate Conflict (UCDP)$_{t-1}$</td>
<td>2.053 (1.635)</td>
<td></td>
</tr>
<tr>
<td>Terrorist Attacks$_{t-1}$</td>
<td>1.005*** (0.007)</td>
<td>0.976*** (0.008)</td>
</tr>
<tr>
<td>Constant</td>
<td>−6.696 (4.457)</td>
<td>−44.455* (18.353)</td>
</tr>
<tr>
<td>Obs.</td>
<td>2,744</td>
<td>2,591</td>
</tr>
<tr>
<td>Wald χ²</td>
<td>22050.16***</td>
<td>20387.58***</td>
</tr>
<tr>
<td>Model</td>
<td>2SLS</td>
<td>2SLS</td>
</tr>
<tr>
<td>Instrumented Instrument</td>
<td>Hate Speech by Parties Weak anti-Defame/Hate Speech Leg.</td>
<td>Hate Speech by Parties Weak anti-Defame/Hate Speech Leg.</td>
</tr>
<tr>
<td>Durbin χ²</td>
<td>3.963**</td>
<td>3.022*</td>
</tr>
<tr>
<td>Wu-Hausman F</td>
<td>3.963**</td>
<td>3.0127*</td>
</tr>
</tbody>
</table>

* p ≤ .1 ** p ≤ .01 *** p ≤ .000.
added to the model, hate speech becomes not significant.\textsuperscript{34} This is a preliminary sign of mediation. The mediation tests statistics are reported below Figure 3 and they confirm that the impact of hate speech on terrorism is mediated through increased political polarization.\textsuperscript{35} Moreover, the percentage of the effect of hate speech on terrorism that is mediated through increased polarization is quite high: around 84 percent.\textsuperscript{36} I also conducted a sensitivity analysis to determine whether the sequential ignorability assumption is problematic in the mediation analysis. To do this, I used a technique developed by Hicks et al. (2011) that calculates the threshold for correlation between the error terms of the relationships between hate speech and polarization ($\varepsilon_{i2}$) and polarization and terrorism ($\varepsilon_{i3}$) that would indicate a violation of the sequential ignorability assumption. Using this technique\textsuperscript{37} I calculated the threshold to be .3.\textsuperscript{38} I then calculated the residuals for the relationships between $\varepsilon_{i2}$ and $\varepsilon_{i3}$ and then tested them for correlation, finding

\begin{table}[h]
\centering
\begin{tabular}{lccc}
\hline
 & \textbf{Mean} & \textbf{95\% Confidence Interval} \\
\hline
ACME1: & .0726536 & .0619507 & .0835709 \\
ACME0: & .0712908 & .0592537 & .0836453 \\
Direct Effect1: & .0144435 & -.0107926 & .0390475 \\
Direct Effect0: & .0130806 & -.0098971 & .0351903 \\
Total Effect: & .0857342 & .0679763 & .1024302 \\
Average Mediation: & .0719722 & .0606182 & .0835829 \\
Average Direct Effect: & .013762 & -.0103449 & .0369739 \\
Percent of Total Effect Mediated: & .839411 & .7026465 & 1.058791 \\
\hline
\end{tabular}
\end{table}

\textbf{Figure 3.} Test of mediation.

\textsuperscript{34}$\beta = .064$ presented not in parentheses in Figure 3.

\textsuperscript{35}Average mediation = .071***.

\textsuperscript{36}Percentage of Total Effect Mediated = .839411.

\textsuperscript{37}Stata command “medsens”.

\textsuperscript{38}Rho at which ACME = 0.
them to be correlated at .0221.\textsuperscript{39} This is below the threshold, suggesting, but not confirming, that the sequential ignorability assumption is not problematic in the mediation analysis.

**Conclusion**

In summary, I find that hate speech by politicians is a statistically significant, and substantive, driver of domestic terrorism in countries. When politicians frequently use hate speech in their public statements, the rate of domestic terrorism increases by almost nine-fold, considering that countries where hate speech is a rare occurrence experience around one attack per year while countries where hate speech is common experience between nine and ten attacks per year, all things being equal. I also find that hate speech boosts domestic terrorism by stoking political polarization that makes domestic terrorism more likely, shedding some light on a causal mechanism linking hate speech with terrorism. The principal conclusion to be drawn from these findings is that politicians’ words have security consequences. The findings also suggest that increased civility on the part of political figures, or projection of tolerance in politicians’ public statements, could yield a substantial security dividend by reducing domestic terrorism.

The study also leaves a number of questions that could be answered in future investigations. For example, what type of terrorism does hate speech motivate? Does hate speech have a different, or particularly strong, impact on lone actor (“lone wolf”) terrorists who might be more susceptible to the radicalizing effects of politician speech than formal, organized armed group that use terrorism as part of a long-term strategy?\textsuperscript{40} Is the relationship between hate speech by political figures and increased terrorism evident at the individual level? Future work could determine, through an experimental research design, whether subjects who are exposed to samples of hate speech by politicians are more likely to endorse, support or sympathize with terrorism or actors that use terrorism. Such a study could also investigate the mediating role that increased partisan polarization – attitudes toward individuals who are supporters of rival political parties or movements – plays in the relationship between hate speech and terrorism. Does hyperbolic, partisan media or social media exacerbate the effect of political hate speech on polarization and terrorism? Future research could examine the role that different types of media or forms of communication or social media technologies might play in amplifying the relationship between politician hate speech, polarization and

\textsuperscript{39}All of these results are available from the author.

\textsuperscript{40}Unfortunately, scholars lack reliable cross-national, time-series data on lone perpetrators. The latest iteration of the GTD 2019 does include a variable that identifies attacks by individuals unaffiliated with terrorist groups (“individual”). However, the GTD codebook includes a disclaimer stating that the individual variable reliably identifies only a portion of lone perpetrators and is affected by significant missingness. Preliminary descriptive statistics suggests that data missingness for the “individual” variable is not random. For example, there are only 223 lone perpetrator attacks in the GTD (comprising around 2.14 percent of attacks since 2000) and around 21.5 percentage of these occurred in just one country: the United States. This suggests to me that due to coding challenges and missingness, the distribution of lone perpetrator attacks in GTD is highly skewed.
domestic terrorism. Does hate speech by political figures not only increase the frequency of domestic terrorism in countries, but also the intensity? Follow up studies could examine the impact of hate speech and polarization on terrorist attack casualties or on the prevalence of more extreme types of terrorism or terrorist movement target selection. Finally, does political hate speech and polarization contribute to other negative outcomes, such as political regime instability? Though this paper focuses on the impact of hate speech by politicians on terrorism, it is possible that hate speech, and resulting political polarization, has more profound effects on fundamental political order within societies.

References


