Making America Hate Again? Twitter and Hate Crime under Trump

Karsten Müller* and Carlo Schwarz†

May 14, 2018

Abstract

Social media has come under increasing scrutiny for reinforcing people's pre-existing viewpoints which, it is argued, can create information "echo chambers." We investigate whether social media motivates real-life action, with a focus on hate crimes in the United States. We show that the rise in anti-Muslim hate crimes since Donald Trump's presidential campaign has been concentrated in counties with high Twitter usage. Consistent with a role for social media, Trump's Tweets on Islam-related topics are highly correlated with anti-Muslim hate crime after, but not before the start of his presidential campaign, and are uncorrelated with other types of hate crimes. These patterns stand out in historical comparison: counties with many Twitter users today did not consistently experience more anti-Muslim hate crimes during previous presidencies.

*Warwick Business School, University of Warwick, k.muller@warwick.ac.uk

†Department of Economics, University of Warwick, Centre for Competitive Advantage in the Global Economy (CAGE), c.r.schwarz@warwick.ac.uk

Introduction

President Donald Trump is arguably the most prominent politician to actively use social media today. Many commentators have criticized Trump for his inflammatory rhetoric on Twitter and cited Trump as an example of how social media can increase political polarization (New York Times, 2017). The empirical evidence on this topic, however, is extremely limited. While polarization has increased in recent years (Fiorina and Abrams, 2008; Gentzkow, 2016), existing studies have found no or even a negative correlation with social media (Boxell et al., 2017; Barberá, 2014).[1]

In this paper, we investigate the role of social media in the context of hate crimes, which the Federal Bureau of Investigation (FBI) defines as "crimes in which the perpetrators acted based on a bias against the victims race, color, religion, or national origin" (FBI, 2016 (accessed March 3rd, 2018). One interpretation of such crimes is that they are particularly extreme actions that follow from polarization. Could it be that the limited role for social media on *average* polarization in previous studies hides an effect on extreme outcomes? In particular, one might expect that the prejudices at the heart of hate crimes could be re-enforced by social "echo chambers"(Sunstein, 2000, 2002, 2017) and flare up spontaneously when triggered by political events (Fouka and Voth, 2013).

To start, we document that the weekly number of anti-Muslim hate crimes, as recorded by the FBI, has increased during Donald Trump's presidency. Indeed, the number of weekly hate crimes committed against Muslims under Trump is twice as high as under Obama and 50% higher than under Bush. This is particularly striking because Bush's term included a temporary 10-fold increase in anti-Muslim hate crimes following the 9/11 terror attacks – the largest spike since the beginning of the FBI records in 1990 (Gould and Klor, 2016; Panagopoulos, 2006; Hanes and Machin, 2014).

We show that the increase in hate crimes targeting Muslims begins with the start of Donald Trump's presidential campaign and is almost exclusively driven by counties with high Twitter usage. The shift in anti-Muslim hate crimes does not seem to be driven by differences in local voting patterns, ethnic composition, economic performance, or general crime rates across counties. In fact, counties with many Twitter users are considerably less likely to vote Republican (and Trump in particular), have a lower share of Whites, lower poverty among children and the elderly, and are less likely to be in rural areas. Moreover, the largest spikes in hate crime coincide with key events during Trump's presidential campaign,

[1] A separate literature has analyzed the effects of the media on elections and other political outcomes. See, among others, the work by Adena et al. (2015), DellaVigna et al. (2014), Stephens-Davidowitz (2014), Gavazza et al. (2015), Gentzkow (2016), and Martin and Yurukoglu (2017). Enikolopov et al. (2011), Enikolopov et al. (2016), and Müller and Schwarz (2018) study the role of social media.

in particular his election as the 45th president of the United States, call for a Muslim ban, and reaction to the Brussel terror attacks.

These correlations, however, may not isolate a pure “Trump effect”, because it is unclear to which extent counties with many and few Twitter users differ in unobservable dimensions we are unable to measure. For example, it is possible that areas where a large share of the population uses Twitter today are generally more hostile towards Muslims, particularly when triggered by an event such as Trump's presidency. A look further back in time reveals that the spike in anti-Muslim hate crimes in counties with high Twitter usage under Trump is unprecedented compared to previous presidencies since 1990. If we were merely capturing persistent local differences in other factors, we would expect hate crimes to be systematically more prevalent in these counties during earlier periods as well. In the data, however, anti-Muslim hate crimes have only become highly concentrated in areas with many Twitter users under Trump. This even holds true compared to the presidencies of Barrack Obama, whose election was widely regarded as a critical moment for inter-group relations in the United States at the time – and during which Twitter use was already widespread. It also holds true for the first presidency of George W. Bush, and thus the period including 9/11.

Finally, we analyze Trump's Twitter feed and show a highly robust time-series correlation between his tweets on Islam-related topics and the number of anti-Muslim hate crimes after his campaign start. Consistent with our findings on the shift in the geography of hate crimes, this effect is particularly strong in counties where many people use Twitter. There is no correlation between Trump's tweets and hate crimes with other motives, which suggests that we are not merely capturing waves of anti-minority sentiment. Importantly, we also find no such link for the period before Trump's presidential campaign.

Taken at face value, this evidence suggests that, with the start of Donald Trump's presidential campaign, social media may have come to play a role in the propagation of hate crimes – and thus political polarization. It is important to stress that we do not claim that Donald Trump's presidential campaign, or social media itself, is the immediate *cause* of the increase in anti-Muslim hate crime. Rather, we document that Trump's presidential campaign marks a cesura for the connection between Twitter usage and hate crimes in the United States. Our findings are thus consistent with the interpretation that Trump's presidential campaign aided an unraveling of social norms that made people more willing to express views that were previously deemed socially unacceptable (Bursztyn et al., 2017).

Our results suggest that this unraveling may have prompted a few people to carry out hate crimes, and that social media – and Twitter in particular – may be non-negligible for understanding how hateful sentiment spreads since Trump's presidential campaign. This is supported by previous evidence that Twitter users are predominately male and more ideologically extreme than the average population (Barberá and Rivero, 2015).

Data on Hate Crime and Twitter

We draw on five primary data sources for our analysis (see Supplementary Material 1 for more details). First, we use the FBI's hate crime data for the years 1990 until 2016. The data set contains all hate crimes in the US that are reported to the FBI, including the exact date of the crime. The original reporting agency (ORI) also makes it possible to assign hate crimes to counties. The FBI also differentiates hate crimes by the motivating bias (e.g. anti-Muslim). These data are well-known to drastically under-report the actual number of hate crimes (ProPublica, 2017 (accessed March 3rd, 2018; NBC News, 2017 (accessed March 3rd, 2018); however, they have the advantage that they contain the exact date and location of each crime, so that we can map them to counties and weeks.[2]

Secondly, we create a measure of Twitter usage in each US county using two data sets of geo-located Tweets from Zubiaga et al. (2016) and Follow the Hashtag (2016 (accessed March 3rd, 2018). When combined, these sources contain over 1.2 million tweets collected in the years 2014, 2015, and 2016, which we use to identify the Twitter usage for each county.

Third, we collect the universe of tweets from Donald Trump from Brown (2018 (accessed March 3rd, 2018). Overall, this dataset contains 32,632 tweets for the time period of April 2009 until December 2016.

Fourth, we obtain additional demographic control variables at the county level from the United States Census. Among others, the data contain the total population for each county for the years 1990 through 2016. Further information on electoral outcomes as well additional indicators were collected by Kirkegaard (2016). The data are were downloaded from OpenDataSoft.com.

Last, we use data from the Media Consumption Survey by PEW research to investigate the correlation between Twitter usage and general media consumption.

Presidents and Trends in Anti-Muslim Hate Crime

We begin by investigating whether there has been an increase of anti-Muslim hate crimes since the election of Donald Trump. To assess this possibility, we compare the weekly average and

[2]Note that reporting bias is unlikely to drive our results. First, the US-wide trend of hate crimes reported to the FBI is likely to be highly correlated with the "true" hate crimes trend. Second, we accommodate potential geographical reporting differences in our cross-sectional tests by, implicitly, subtracting the mean number of crimes within (1) a county, as well as (2) a state in a given week, using fixed effects panel regressions.

median number of hate crimes depending on which president is in charge in Figure 1, which also plots the 95% confidence interval around the mean. Because the FBI only publishes hate crime data for the previous year in November, we only have information until December 31, 2016, i.e. the first few weeks after Trump's election. Hence, we compare hate crimes under Trump to two comparison groups: (1) the average number of hate crimes in each previous presidency since 1990 in panel (a); and (2) hate crimes during the period between each president's election in November until the end of the year in panel (b), which gives us approximately equal intervals for each president. Restricting the sample to the year-end weeks makes it unlikely that seasonal factors influencing hate crimes might lead to false conclusions about their frequency under Trump.

Over the 26-year period for which the FBI publishes data, the number of hate crimes against Muslims in the United States has increased. Anti-Muslim hate crimes were somewhat less common under Obama than under Bush. Most strikingly, the post-election period of Trump is a clear outlier by historical standards, particularly when comparing post-election weeks. However, it is striking that the increase still persists in comparison to George W. Bush's entire first term, which included the largest recorded spike in anti-Muslim hate crimes in the wake of 9/11 (Gould and Klor, 2016; Panagopoulos, 2006; Hanes and Machin, 2014).

(a) Full Presidency **(b) Equal Periods**

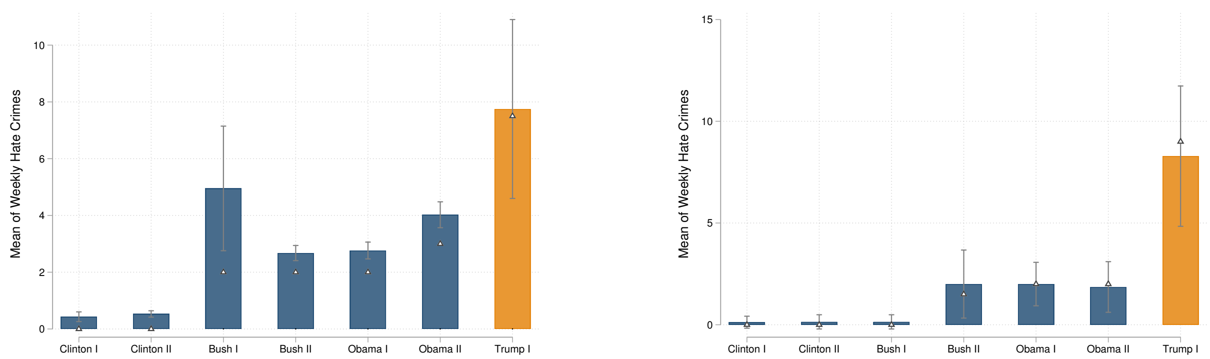


Figure 1: Average weekly anti-Muslim hate crimes in the United States since 1990, by president: This figure plots the average weekly number of hate crimes, by president. Panel (a) shows the full period of the presidency. Panel (b) shows the bar graphs for the identical end-of-year periods after the election. The bars indicate the 95% confidence intervals. The triangle marks the median.

Taken together, we conclude that up to the end of 2016 the frequency of anti-Muslim hate crimes was higher under Donald Trump than previous presidents. It is important to point out that this increase specific to hate crimes against Muslims: In Supplementary Material 2 we show that there is no clear trend in the total number of hate crimes. We also find that the beginning of Donald Trump's *presidential campaign* is an important demarcation point in the number of anti-Muslim crimes, which we will explore in more detail below. Next, we turn to the question whether hate crime has become more responsive to social media activity, in particular Twitter.

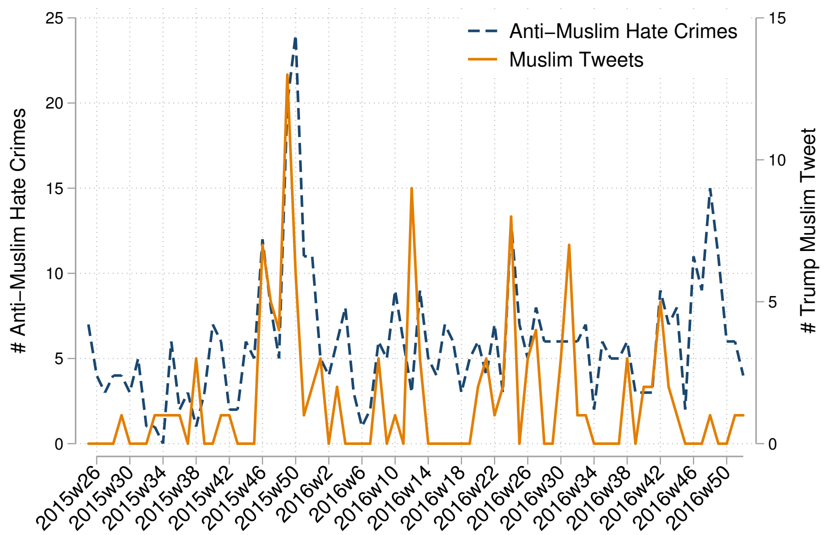
Twitter and Hate Crime: Is Trump Different?

President Trump has been using Twitter frequently since joining the messenger service in March 2009, racking up a total number of more than 32,600 tweets. While he only had around 2.7 million followers in December 2014 (BBC, 2016 (accessed March 3rd, 2018), the number is above 41 million today. Many of his incendiary or derogatory comments, in particular, are often further propagated through news articles reporting on his statements. To identify changes in his reach, Figure A.2 in the Supplementary Materials 1 shows the monthly of re-tweets that his own tweets received for the period up to the end of 2016. Trump's presidential campaign start, indicated with the vertical line, marks a substantial shift: starting in June 2015, the number of re-tweets increased steadily, dwarfing the attention he had received previously. This shows that Trump did not only accumulate more followers, but that his tweets were also much more widely shared.

As a next step, we investigate whether Trump's tweets are correlated with real-life actions. In particular, we exploit weekly variation in the content of his tweets and link them to US-wide hate crimes against Muslims in a simple time-series framework. The Supplementary Material 3 provides descriptive statistics and compares both hate crimes and Trump tweets before and after the campaign start. Panel (a) in Figure 2a plots the number of Trump tweets on Muslim-related topics in a week and the number of anti-Muslim hate crimes committed (see Supplementary Material 1 for details on the classification of tweets). The two series show substantial co-movement, suggesting that Trump's tweets are highly correlated with the number of hate crimes against Muslims.

In Table 1 Panel (B) we show that this relationship is also significant in a statistical sense: Trump's tweets about Muslims have substantial forecasting ability for anti-Muslim hate crimes until about 3 or 4 weeks into the future. Intuitively, this predictive ability decays almost monotonically with the forecasting horizon. Strikingly, Trump's tweets have zero predictive ability for the period before his presidential campaign (see Table 1 Panel (A)). This is particularly noteworthy because the 318 weeks between Trump joining Twitter and his campaign announcement allow for much more statistical power than the 80 post-announcement weeks. These differences are also statistically significant, as indicated by the p-values from a χ^2-test for the equality of coefficients for the models in these separate samples in the bottom row. Trump's Muslim tweets alone predict more than 20% of the variation in

(a) **Trump Tweets and Hate Crime against Muslims**



(b) **Trump tweets on Islam-related topics and hate crimes, by group**

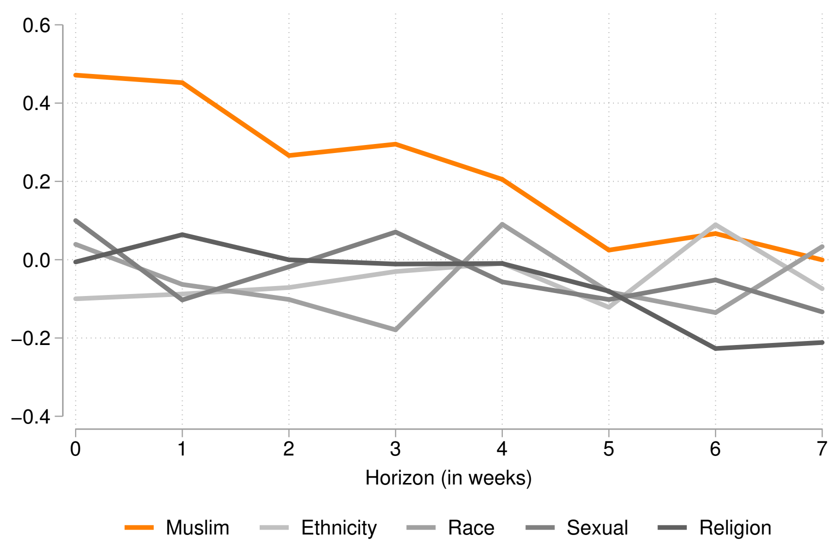


Figure 2: Panel (a) shows the weekly number of Donald Trump's Islam-related tweets and the number of anti-Muslim hate crimes in the US after the start of Trump's presidential campaign. Panel (b) plots the correlation of hate crimes against different groups and Trump's tweets about topics related to Muslims at horizons between 0 and 7 weeks. The sample period are the weeks after Donald Trump announced his presidential campaign (the week of June 16, 2015).

anti-Muslim hate crimes in the same week, but only after his campaign start; the explanatory power is less than 1% before. In the Supplementary Material 3, we conduct a battery of additional exercises and find that these associations are remarkably robust.

Do Trump's tweets simply reflect *general* waves of anti-minority sentiment? As a simple counterfactual test, we next investigate whether Trump's messages about Muslims are also correlated with hate crimes against other minorities, in particular those motivated by ethnicity, race, sexual orientation, or religions other than Islam. Panel (b) in Figure 2b plots the predictive ability of Trump's tweets about Islam-related topics for different types of hate crimes, including those against Muslims, obtained from our time series regression framework. Clearly, Trump's tweets are only correlated with crimes against the Muslim minority, not other hate crimes. This suggests that we are not merely capturing anti-minority sentiment, but a Muslim-specific time trend.

In additional exercises, reported in Supplementary Material 6, we find complementary results for Hispanics. While the association weakens somewhat in the immediate run-up to the election in mid-2016, Table A.10 and Table A.11 show that Trump's tweets about Hispanics have considerable predictive power for Ethnicity-based hate crimes. Again, this only holds true for the period after his campaign start, and not for other types of crime biases.

These correlations are consistent with at least two possible explanations. One is that Donald Trump became more reactive to US-wide anti-Muslim sentiments. Another possibility is that Trump's tweets themselves contribute to a climate that enables hate crimes. Clearly, we cannot disentangle these possibilities using the correlations above. To shed further light on this, we next ask whether the increase of anti-Muslim hate crimes under Trump shown in Figure 1, and the correlation with his Twitter posts, are concentrated in areas with many Twitter users.

County-Level Evidence

To better understand the strong time-series correlation between Trump's tweets and hate crimes against Muslims, we investigate the geography of Twitter usage in the United States. In the Supplementary Material 1, we show that the number of Twitter users per capita varies considerably, even across neighboring counties. In Figure 3a, we plot the average monthly number of anti-Muslim hate crimes per capita in counties above and below the 90th percentile in Twitter usage per capita over time, where the vertical line again marks the beginning of Trump's presidential campaign. (In Supplementary Material 4, we show that the results are very similar if we use other percentiles as cutoffs instead.) To make the numbers more easily comparable, we have standardized the number of hate crimes per capita to have zero mean

Table 1: Time Series Evidence: Trump Tweets and Anti-Muslim Hate Crimes

Dependent Variable:	Anti-Muslim Hate Crimes in							
	t	t+1	t+2	t+3	t+4	t+5	t+6	t+7
Panel A: Before Campaign Announcement								
Muslim Trump Tweet	0.089	0.031	-0.016	0.054	0.052	0.061	-0.013	-0.027
	(0.071)	(0.071)	(0.066)	(0.043)	(0.068)	(0.049)	(0.065)	(0.045)
Observations	318	318	318	318	318	318	318	318
Muslim Tweet Partial R^2	0.008	0.001	0.000	0.003	0.003	0.004	0.000	0.001
Adj. R^2	0.029	0.017	0.014	0.013	0.008	0.014	0.001	0.007
Panel B: After Campaign Announcement								
Muslim Trump Tweet	0.471***	0.452**	0.266**	0.295*	0.205	0.024	0.067	-0.000
	(0.153)	(0.178)	(0.122)	(0.170)	(0.133)	(0.071)	(0.097)	(0.091)
Observations	80	79	78	77	76	75	74	73
Muslim Tweet Partial R^2	0.221	0.204	0.071	0.087	0.042	0.001	0.004	0.000
Adj. R^2	0.433	0.270	0.105	0.104	0.057	0.009	0.033	0.080
Equal coefficients (χ^2)?	0.018	0.022	0.021	0.111	0.377	0.712	0.477	0.755

Notes: This table presents the estimated coefficients from an Ordinary Least Squares regression of weekly hate crimes at the horizon indicated in the top row on tweets by Donald Trump containing Muslim-related words. All variables are standardized with mean 0 and standard deviation 1 and thus directly comparable across the regressions. In Panel A, we only include the period before Donald Trump announced his presidential campaign (the week of June 16, 2015). Panel B includes the period after his announcement until the end of the available hate crime data on December 31, 2016. All regressions include one lag of the dependent variable, as suggested by a lag selection test based on the Bayesian information criterion (BIC). The Partial R^2 is calculated as the difference of the regression model with and without the Muslim Trump Tweet variable. Newey-West standard errors are reported in parentheses, allowing for 6 and 4 lags in Panel A and B, respectively, in line with the lag selection procedure in Andrews and Monahan (1992). The "Equal coefficients" row reports a χ^2 test for the equality of the estimated coefficients in Panel A and B. ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.1 level, respectively.

and unit variance and partialed out county and week fixed effects. This means all values should be interpreted relative to the average in a particular county and week.

While the number of anti-Muslim hate crimes is more volatile in counties with high Twitter usage (in orange), we can see that the number of hate crimes was more or less constant since 2009. With the start of Donald Trump's presidential campaign on June, 16th 2015, however, we observe a disproportional increase in the number of hate crimes in those counties where many people use Twitter. There is no comparable increase in counties with low twitter usage (in blue). In contrast, Figure 3b shows that there is no such differential effect for all other hate crimes taken together. As we show in Supplementary Material 4, these trends are visible independent of the cut-off chosen for high Twitter usage.

The results indicate that the increase in anti-Muslim hate crime is predominantly driven by counties that have a high Twitter usage, which only materialized after the start of Donald Trump's presidential campaign. Interestingly, the three months with the highest number of anti-Muslim hate crimes after Trump's presidential run in Figure 3a coincide with three important events in Trump's campaign, namely (1) his demand for a Muslim travel ban, (2) the Brussel Terror Attacks, in the wake of which Trump endorsed torturing potential terror suspects, and (3) the month of his presidential election.

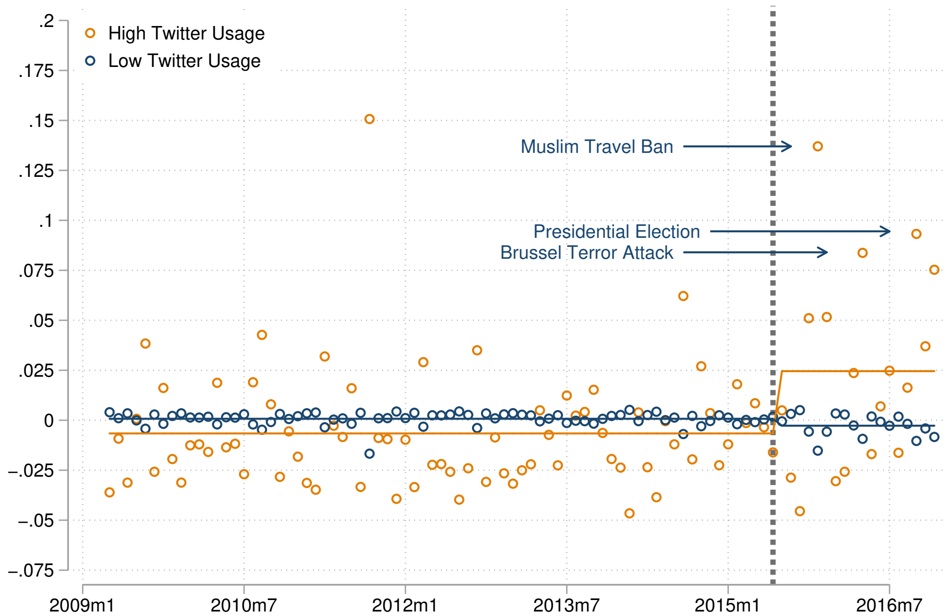
Is the shift in anti-Muslim hate crimes also statistically significant? And can it be explained by observable differences across counties, e.g. differences in voting patterns, ethnic composition, economic conditions, or crime? We investigate these questions by re-examining the evidence presented in Figure 3a more formally in a panel regression framework. The advantage of panel regressions is that we can abstract from other county-level factors that may explain the shift in hate crimes. In particular, we estimate the following equation:

$$\begin{aligned} Hate\,Crime/Pop._{csw} = {} & \beta Pres\,Run_w \times I[90th\,Pct.\,Twitter\,Usage]_{cs} \\ & + \gamma Pres\,Run_w \times Controls_{cs} + \alpha_c + \alpha_{sw} + \epsilon_{csw}, \end{aligned}$$

where $HateCrimes/Pop._{csw}$ indexes the number of anti-Muslim hate crimes per capita in county c in state s in week w. $Pres\,Run_w$ is a dummy variable that is 1 for weeks after Trumps presidential run, while the dummy variable $I[90th\,Pct.\,Twitter\,Usage]_{cs}$ tags counties above the 90th percentile in Twitter usage per capita (and is 0 otherwise). The vector $Pres\,Run_w \times Controls_{cs}$ contains different sets of county-level control variables, which are also interacted with Trump's presidential run dummy (see Table A.3 in Supplementary Material 1 for the full list of controls).

To control for any permanent differences in anti-Muslim hate crimes across counties,

(a) **Anti-Muslim Hate Crimes**



(b) **All other Hate Crimes**

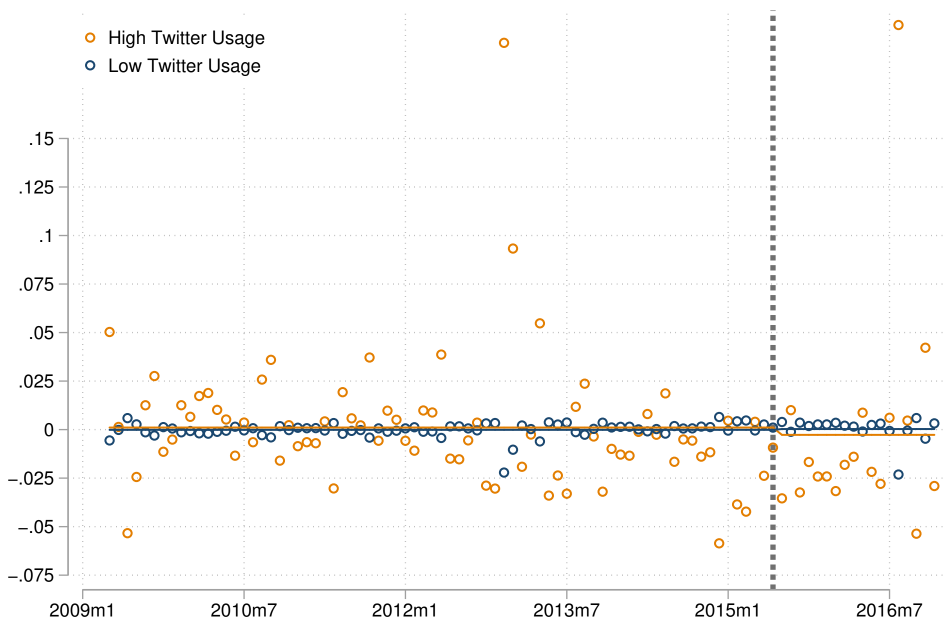


Figure 3: Monthly Hate Crimes by Twitter Usage: This figure plots the monthly number of hate crimes per capita for counties above and below the 90th percentile of Twitter usage. Panel (a) plots anti-Muslim hate crimes and panel (b) all other hate crimes. The vertical line marks the beginning of Trump's presidential campaign.

for example due to higher xenophobia or reporting differences in the FBI data, the regression contains a full set of county fixed effects (α_c). We further control for any shocks that might increase the number of anti-Muslim hate crimes nationwide or in any particular state, e.g. during weeks of Republican primaries in that state, by including state-times-week fixed effects (α_{sw}).[3]

Hence, β measures the relative increase of anti-Muslim hate crimes in counties with high Twitter usage with the start of Trumps presidential campaign. The regression results, presented in Table 2, show that Trump's campaign let to a 3.9% of a standard deviation increase in the number of anti-Muslim hate crimes in counties with high Twitter usage. To test whether this result can be explained by other factors, we next include additional county-level factors, interacted with the $Pres\ Run_w$ variable.

Table 2: Changes in Anti-Muslim Hate Crime with Trump's Presidential Run

Dependent Variable	Anti-Muslim Hate Crime					
	(1)	(2)	(3)	(4)	(5)	(6)
$Pres\ Run \times$	0.039**	0.037**	0.032*	0.031**	0.031**	0.040***
$I[90th\ Pct.\ Twitter\ Usage]$	(0.017)	(0.018)	(0.017)	(0.014)	(0.015)	(0.014)
Voting Controls		Yes				
Ethnicity Controls			Yes			
Demographic Controls				Yes		
Economic Controls					Yes	
Crime Controls						Yes
State-Week FE	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,253,259	1,241,289	1,253,259	1,253,259	1,250,865	483,189
Number of counties	3,141	3,111	3,141	3,141	3,135	1,211
R-squared	0.022	0.022	0.022	0.022	0.022	0.063

Notes: This table presents the estimated coefficients from a regression of hate crimes on Twitter usage described in the text. The dependent variable is the number of hate crimes committed against Muslims per capita. The dependent variable is standardized with mean 0 and standard deviation 1. *Pres Run* is an indicator variable for the time period from the beginning of Donald Trump's presidential run. $I[90th\ Pct.\ Twitter\ Usage]$ is an indicator variable that is 1 for counties above the 90th percentile of Twitter usage per capita. All regression include county and state-week fixed effects. See text for an explanation of the control variables. Robust standard errors in all specifications are clustered by county. ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.1 level, respectively.

First, we test whether our result can be explained by the political affiliation in column (2), which we attempt to measure using the a dummy variable for counties won by Republicans in the 2008, 2012 and 2016 presidential elections. We find no indication that our findings can be explained by differences in the support for the Republican party. In fact, Table A.3

[3] The results are highly similar if we replace state-week with week fixed effects.

in Supplementary Material 1 shows that counties with many Twitter users are considerably *less* likely to vote Republican, especially in the 2016 election. Next, we control for the ethnic composition of counties by including the share of Muslims, Whites, Blacks, Hispanics, and Asians in column (3). Note that Twitter usage is negatively correlated with the share of Whites and positively with that of minorities, including Muslims; conditioning on ethnicity, however, does not affect our findings. Including a host of demographic variables to control more thoroughly for differences in population, gender and age in column (4) also has no effect.

To test whether the effect is driven by poorer counties, we include a number of economic indicators in column (5). Finally, we also include controls for the crime rate (which are only available for a smaller number of counties). Including these variables has virtually no bearing on our results. In Supplementary Material 4 we also test more formally for preexisting trends in our regression; the results suggest that the shift in hate crimes only materialized after Trump's presidential run.

While not conclusive, these findings suggest that the increase in anti-Muslim hate crime with Trump's campaign start in areas with high Twitter usage cannot be easily explained by other observable county differences. Indeed, counties with many Twitter users tend to be urban areas with considerable ethnic minorities and lower poverty that are likely to vote Democratic; this means that these potentially confounding factors are unlikely to explain our results. Survey evidence from the Pew Research Center, presented in Table A.13 in Supplementary Material 8, further suggests that we are not capturing a "Fox News effect" (DellaVigna and Kaplan, 2007; Martin and Yurukoglu, 2017, e.g.): while Twitter users are more likely to get their news from Twitter, they prefer MSNBC and CNN over Fox News.

As a further test, we analyze hate crimes against other minority groups. To do so we estimate the same panel regression as before, but instead use hate crimes against ethnic, racial, sexual, or other religious minorities as an outcome. The results of these regressions are presented in Supplementary Material 4. We do not find a significant increase in the number of hate crimes against these other groups in counties with high Twitter usage after the start of Trump's presidential campaign. However, we find somewhat weaker evidence of a disproportionate increase in hate crimes on Ethnic grounds (which includes anti-Hispanic hate crimes). This meshes well with further time series evidence we present in Supplementary Material 6 on the link between Trump's tweets about Hispanic-related topics and ethnic hate crimes. It also shows that the observed correlations are specific to hate crimes against those minorities arguably most directly verbally targeted by Trump, rather than due to a general up-tick in anti-minority sentiment in counties with high Twitter usage.

Are counties with many Twitter users also those that drive the correlation between

Trump's tweets about Muslims and hate crimes? We investigate this question by re-running the time series regressions presented above separately for counties with high and low Twitter usage in Table A.7 in Supplementary Material 3. When we compare the predictive ability of Trump's twitter posts about Muslims in these samples, the overwhelming predictive ability stems from counties with many Twitter users. This is true even though we hold constant the average number and volatility of hate crimes in these counties.

Quantitatively, the Muslim tweets explain between 13.5 and around 16% of the weekly variation in anti-Muslim attacks in counties with high Twitter usage, depending on whether we use the median, top quartile, or top decile in the Twitter user to population ratio. The tweets only explain between 0 and 10% where fewer people use Twitter. Taken at face value, this further points to a potential role for social media in enabling hate crimes against Muslims in the United States. In the next section, we assess whether the correlations we have established thus far reflect a geographical divide in where hate crimes occur.

Are Areas with High Twitter Usage More Prone to Hate Crimes?

A potential concern with the panel tests we present above may be that counties in which many people use Twitter today may be inherently different in ways that are not captured by the control variables included in our regressions. For example, there may be differences in racial bias (Stephens-Davidowitz, 2014), persistent attitudes towards particular social groups (Voigtlander and Voth, 2012), or historical experiences that might flare up spontaneously when triggered by political events (Fouka and Voth, 2013). To assess the likelihood that such unobservable factors can explain the disproportionate rise of anti-Muslim hate crime in counties with many Twitter users under Trump, we compare the period since Trump's election to all previous presidencies since 1990.

Figure 4 plots the number of anti-Muslim hate crimes split by the indicator for high Twitter usage (as defined by the 90th percentile as before) for all presidents since 1990. Because we want to make sure the findings are not driven by the fact that we only have data on Trump for the few weeks after his election, we again present results based on the total duration of the previous presidencies, as well as the post-election weeks of the election year. To abstract from the fact that such crimes have seen a general upward trend under Trump (see above), we residualize the raw data with *state* $\times$ *week* fixed effects. We also partial out county fixed effects, which further de-means the number of hate crimes by county. As we show in Supplementary Material 5, the results look highly similar without these adjustments.

If our results are driven by the fact that people in areas with many Twitter users generally commit more hate crimes, we should also see more hate crimes occurring there during previous presidencies. The advantage of this approach is that we do not need to pin down the exact differences across counties that may be correlated with Twitter usage to assess whether Trump is different, but instead use the recent history of presidencies as a guide. In comparison to all other presidents, Trump is a striking outlier: the difference in anti-Muslim hate crimes between counties with many and few Twitter users today is considerably higher under Trump compared to other presidents. In fact, the standardized number of per-capita hate crimes under Clinton's two presidencies was *lower* in areas where many people use Twitter today. The differences were relatively muted under Bush, with a slight increase for Obama's two terms; note that, of course, Obama's second term already includes the disproportional increase in hate crimes after Trump's campaign start we document above. The standardized number of crimes since Trump's election, however, show a striking difference of more than 6% of a standard deviation between the two types of counties in per-capita terms.

These patterns are also borne out in the raw data, which we present in Supplementary Material 5. Counties with many Twitter users experienced 12 times more anti-Muslim hate crimes per capita than those with few Twitter users under Trump (or 10 times more in absolute numbers). The corresponding ratios for the previous presidencies since Clinton are 6.7, 1.5, 2.5, 3.2, 3.7, and 3.7 (in chronological order). Note, however, that these raw numbers do not allow us to abstract from US-wide changes in the total number of hate crimes, and are thus less informative for their distribution across counties.

(a) Full Presidency **(b) Equal Periods**

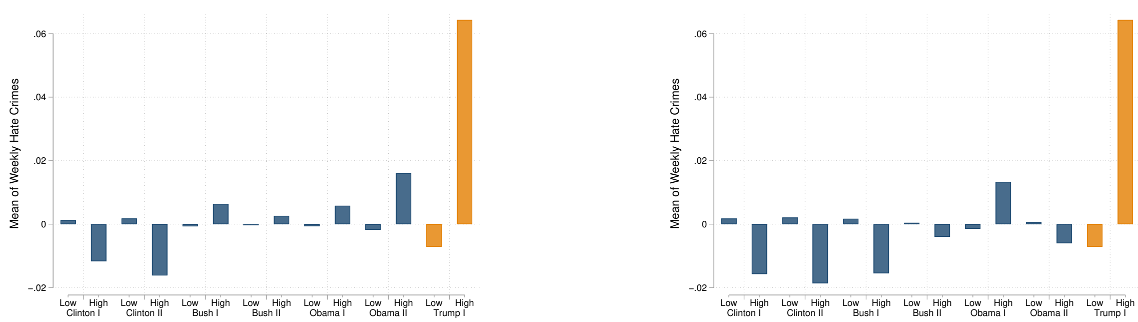


Figure 4: Number of anti-Muslim hate crimes per capita, by presidency and Twitter usage: This figure shows the average weekly number of anti-Muslim hate crimes per capita split by Twitter usage (based on the 90th percentile) and presidency. Panel (a) shows the full period of the presidency. Panel (b) shows the bar graphs for the identical end-of-year periods after the election. Raw data are residualized with county and state-week fixed effects.

Overall, it appears that the concentration of anti-Muslim hate crimes we have documented above is a clear anomaly by historical standards. When comparing hate crimes since Donald Trump's election to those during previous presidencies, we find that they are considerably more common in counties where many people use Twitter, and that these areas were not generally more prone to hate crimes before.

Conclusion

Social media has recently come under scrutiny for its oft-alleged potential to increase citizen polarization by creating informational "echo chambers" (Sunstein, 2009, 2017). Yet, the empirical evidence on this question is limited and has led to widely differing conclusions (Boxell et al., 2017). Consistent with evidence that social media can motivate real-life action (Enikolopov et al., 2016; Müller and Schwarz, 2018), we find that the recent increase in anti-Muslim hate crimes following President Trump's presidential campaign has been concentrated in counties with higher Twitter usage. This is particularly striking given that the total number of hate crimes does not seem to vary much between presidential terms since 1990. Importantly, hate crimes were not concentrated in areas where many people use Twitter today during previous presidencies over the past 26 years. This suggests that we are not capturing a flaring up of long-held political beliefs or historical racial animosity under Trump.

While only suggestive, our findings are consistent with a role for social media in propagating incendiary online comments. In line with this hypothesis, we find that Trump's anti-Muslim tweets are highly correlated with the number of anti-Muslim hate crimes, particularly where Twitter usage is high, but only after the start of his presidential campaign. Going forward, it remains for controlled experiments and different case studies to show how tight the link between social media and real-life outcomes is in other settings.

While we have focused on a particularly negative outcome – hate crimes – social media may well have a positive impact in other areas of life. Further, we would like to caution against using empirical regularities such as the ones documented here as a sole basis for policies with the goal of restricting online communication. History is ripe with cautionary tales of how excessive state power over the media can abed or enable authoritarian rule. The complex trade-offs that policy makers face in this regard thus require nuanced discussion and, above all, more evidence.

References

M. Adena, R. Enikolopov, M. Petrova, V. Santarosa, and E. Zhuravskaya. Radio and the rise of the nazis in prewar germany. *The Quarterly Journal of Economics*, 130(4):1885–1939, 2015.

D. W. K. Andrews and J. C. Monahan. An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. *Econometrica*, 60(4):953–966, 1992. ISSN 00129682, 14680262. URL `http://www.jstor.org/stable/2951574`.

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

P. Barberá. How social media reduces mass political polarization. evidence from germany, spain, and the us. *Job Market Paper, New York University*, 46, 2014.

P. Barberá and G. Rivero. Understanding the political representativeness of twitter users. *Social Science Computer Review*, 33(6):712–729, 2015.

BBC. Trump on twitter: A history of the man and his medium. `http://www.bbc.com/news/world-us-canada-38245530`, 2016 (accessed March 3rd, 2018).

L. Boxell, M. Gentzkow, and J. M. Shapiro. Greater internet use is not associated with faster growth in political polarization among us demographic groups. *Proceedings of the National Academy of Sciences of the United States of America*, page 201706588, 2017.

B. Brown. The trump twitter archive. `http://www.trumptwitterarchive.com/`, 2018 (accessed March 3rd, 2018).

L. Bursztyn, G. Egorov, and S. Fiorin. From extreme to mainstream: How social norms unravel. 2017.

S. DellaVigna and E. Kaplan. The fox news effect: Media bias and voting. *The Quarterly Journal of Economics*, 122(3):1187–1234, 2007.

S. DellaVigna, R. Enikolopov, V. Mironova, M. Petrova, and E. Zhuravskaya. Cross-border media and nationalism: Evidence from serbian radio in croatia. *American Economic Journal: Applied Economics*, 6(3):103–32, July 2014. doi: 10.1257/app.6.3.103.

R. Enikolopov, M. Petrova, and E. Zhuravskaya. Media and political persuasion: Evidence from russia. *The American Economic Review*, 101(7):3253–3285, 2011.

R. Enikolopov, A. Makarin, and M. Petrova. Social media and protest participation: Evidence from russia. 2016.

FBI. Hate crimes. https://www.fbi.gov/investigate/civil-rights/hate-crimes, 2016 (accessed March 3rd, 2018).

M. P. Fiorina and S. J. Abrams. Political polarization in the american public. *Annual Review of Political Science*, 11:563–588, 2008.

Follow the Hashtag. 200,000 usa geolocated tweets. free twitter dataset. http://followthehashtag.com/datasets/free-twitter-dataset-usa-200000-free-usa-tweets/, 2016 (accessed March 3rd, 2018).

V. Fouka and H.-J. Voth. Reprisals remembered: German-greek conflict and car sales during the euro crisis. 2013.

A. Gavazza, M. Nardotto, and T. M. Valletti. Internet and politics: Evidence from uk local elections and local government policies. 2015.

M. Gentzkow. Polarization in 2016. *Toulouse Network of Information Technology white paper*, 2016.

E. D. Gould and E. F. Klor. The long-run effect of 9/11: Terrorism, backlash, and the assimilation of muslim immigrants in the west. *The Economic Journal*, 126(597):2064–2114, 2016.

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

E. O. Kirkegaard. Inequality across us counties: An s factor analysis. *Open Quantitative Sociology & Political Science*, 2016.

G. J. Martin and A. Yurukoglu. Bias in Cable News: Persuasion and Polarization. *American Economic Review*, 107(9):2565–2599, September 2017.

K. Müller and C. Schwarz. Fanning the flames of hate: Social media and hate crime. *Working Paper*, 2018.

NBC News. Advocates warn of possible underreporting in fbi hate crime data , by Chris Fuchs. https://www.nbcnews.com/news/asian-america/advocates-warn-possible-underreporting-fbi-hate-crime-data-n830711, 2017 (accessed March 3rd, 2018).

New York Times. Trump shares inflammatory anti-muslim videos, and britains leader condemns them, by Peter Baker and Eileen sullivan. 2017.

C. Panagopoulos. The polls-trends: Arab and muslim americans and islam in the aftermath of 9/11. *International Journal of Public Opinion Quarterly*, 70(4):608–624, 2006.

ProPublica. Why america fails at gathering hate crime statistics, by Ken Schwencke. `https://www.propublica.org/article/why-america-fails-at-gathering-hate-crime-statistics`, 2017 (accessed March 3rd, 2018).

S. Stephens-Davidowitz. The cost of racial animus on a black candidate: Evidence using google search data. *Journal of Public Economics*, 118:26–40, 2014.

C. R. Sunstein. Deliberative trouble? why groups go to extremes. *The Yale Law Journal*, 110(1):71–119, 2000.

C. R. Sunstein. The law of group polarization. *Journal of Political Philosophy*, 10(2):175–195, 2002.

C. R. Sunstein. *Republic. com 2.0.* Princeton University Press, 2009. ISBN 1400827833.

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

N. Voigtlander and H.-J. Voth. Persecution perpetuated: The medieval origins of anti-semitic violence in nazi germany. *The Quarterly Journal of Economics*, 127(3):1339–1392, 2012.

A. Zubiaga, A. Voss, R. Procter, M. Liakata, B. Wang, and A. Tsakalidis. Tweet geolocation 5m. 4 2016.

A Supplementary Materials: Making America Hate Again

A.1. Supplementary Materials 1: Additional Details on Data

FBI Hate Crime Data

As a measure for the number of hate crimes, we use the FBI's hate crime data for the years 1990 until 2016. This data set contains all hate crimes in the US that have been reported to the FBI. The data report the exact date of the crime, thereby allowing us to construct the time-series trends of hate crimes in the US. The FBI's data also differentiate hate crimes by the motivating bias. Overall the FBI differentiates between 19 bias motivations, divided up into anti-racial, anti-religious, anti-ethnicity, and anti-sexual hate crimes. This allows us to identify hate crimes against Muslims. The other categories used in the paper are defined according to the codes listed in Table A.1.

Table A.1: FBI Hate Crimes Codes

Hate Crime Category	FBI Codes
Anti-Muslim	24
Anti-Ethnic	32, 33
Anti-Racial	11, 12, 13, 14, 15
Anti-Religious	21, 22, 23, 25, 26, 27
Anti-Sexual	41, 42, 43, 44, 45

Notes: This table lists the FBI Bias Codes that were used to define the categories used in the paper.

In total, the data set contains 185,294 reported hate crimes, of which 3,093 targeted Muslims. The FBI data also contain the name and county of the reporting agencies. Using the FIPS county code of the reporting agency, we assign the hate crimes to counties. In the rare cases where an agency is located in more than one county we assign the hate crime to all counties the agency is active in. Appendix A.1. show the geographic distribution of Hate Crimes over the mainland USA (Alaska and Hawaii are not show in map). The FBI hate crime data do not contain information on the US territories of Virgin Island, Puerto Rico, Northern Mariana Islands, American Samoa, and Guam.

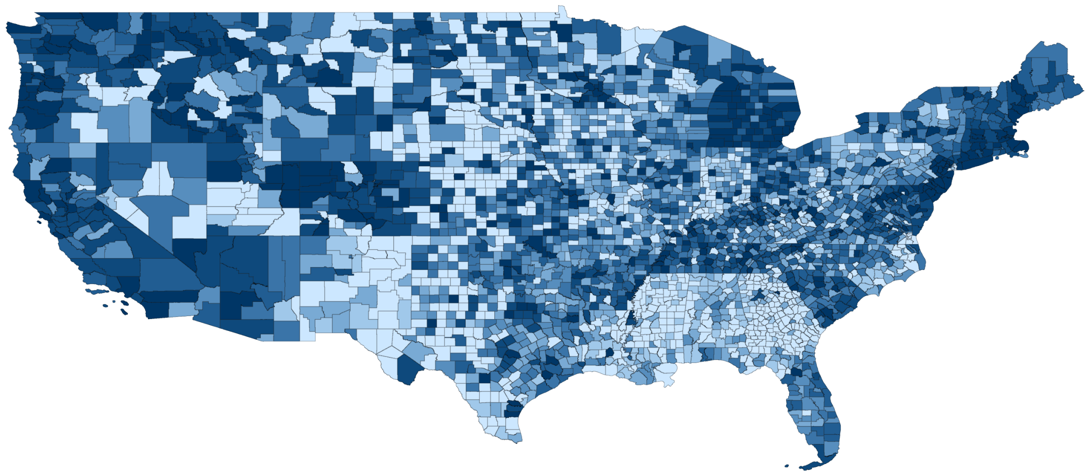


Figure A.1: Hate Crime per Capita by US County: This map show the number of hate crimes per capita in each US counties as reported in the FBI data.

Trump Twitter Data

We downloaded the universe of tweets from Donald Trump from The Trump Twitter Archive. Overall, the Trump Twitter Archive contains 32,632 tweets for the time period of April 2009 until December 2016. The data contains the text of the tweets as well as the number of retweets his tweets received. The text of the tweets allows us to identify tweets of Donald Trump about Muslims or Islam and the number of re-tweets these tweets received. We classify such tweets in a three-step process. First, we use a text search algorithm tagging tweets containing the words "sharia" (1 tweet), "refugee" (22 tweets), "mosque" (3 tweets), "muslim" (60 tweets), "islam" (64 tweets), and "terror" (182 tweets). Second, we read the tweets to verify that we are not picking up unrelated topics by mistake. Third, for each week, we sum the number of Trump tweets containing at least one of these words. In practice, we found that the inclusion or exclusion of single words made little difference; the tweets containing the words "terror", "islam", and "refugee" had the strongest individual effects on contemporaneous hate crimes.[4]

Geocoded Tweet Data

The first source of geolocated Tweets is Zubiaga et al. (2016), which can be downloaded from zubiga.org. The authors collected and geocoded 5 million tweets in October 2014 and October 2015 from all over the world. We keep the 1,061,085 Tweets that are from the United States. The second source for geolocated Tweets is Follow the Hashtag (2016 (accessed March 3rd,

[4]In the baseline time series regression, the estimated coefficients yielded t-statistics of 3.25, 2.7, and 2.16, respectively.

Table A.2: Search Terms for Trump Tweets

Islam-Related Terms	Hispanic-Related Terms
mosque	latino
muslim	hispanic
islam	immigra
terror	illegal alien
sharia	border
refugee	wall

Notes: This table lists the terms that Trump's tweets were searched for to define tweets about Muslims or Hispanics.

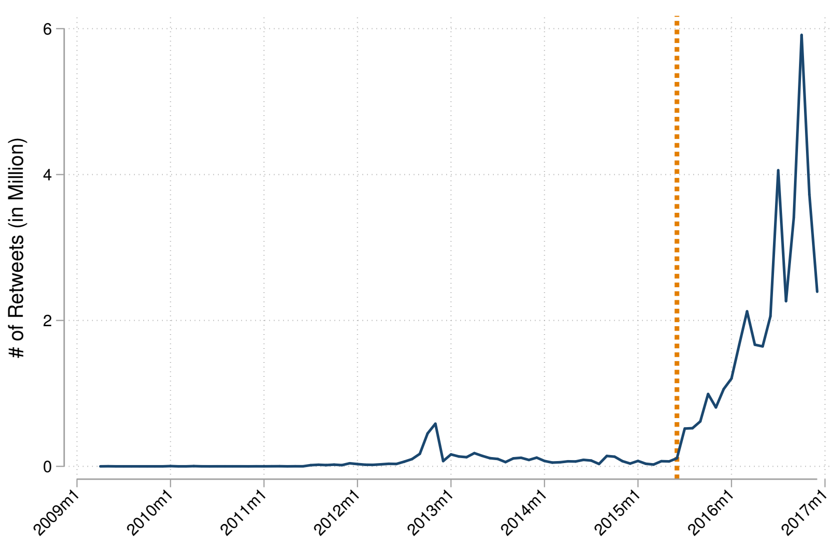


Figure A.2: Trump's Twitter Reach: This figure plots the number of monthly retweets Donald Trump received. The vertical line marks the start of his presidential campaign in June 2015.

2018), from which we obtain 178,507 Tweets that were collected within 48 hours on the 15th and 16th April 2016. This data set is available from followthehashtag.com.

Both data sets contain exact geo-coordinates that allow us to assign each tweet to US counties using the 2017 TIGER/Line Shapefiles provided by the US Census. We create a measure of Twitter usage in each county by dividing the number of Twitter users by county population. Figure A.3 visualizes our Twitter usage measure for the USA (Alaska and Hawaii are not shown).

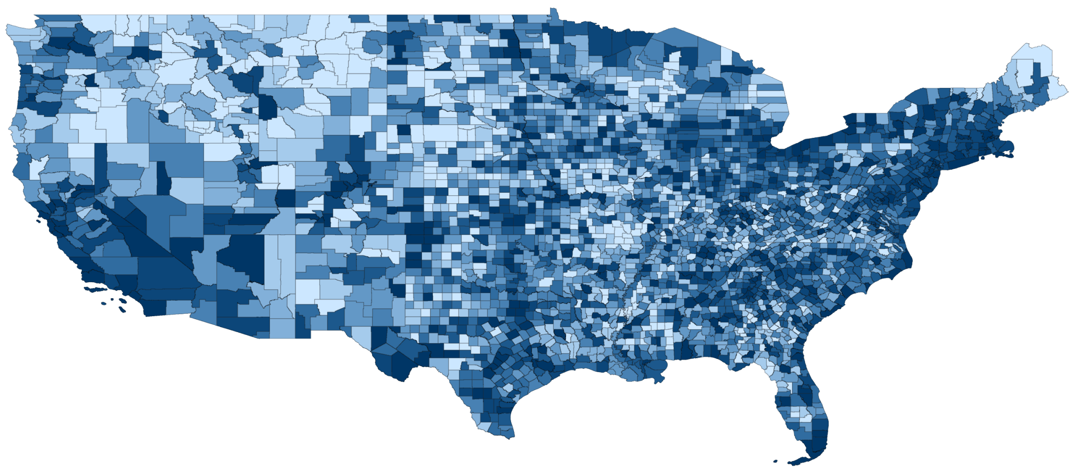


Figure A.3: Twitter Usage per capita by US County: This map plots the Twitter usage measure for in each US y as measured by the number of geo-located tweets divided by population.

Control Variables

Information on the yearly population numbers as well as the gender and age composition for each US county were obtained from the website of the US census. We unify the FIPS county codes for the 7 counties that were split or merged over the 26 year period, so the FIPS county codes align with the 2017 county boundaries. The number of Muslims in each US county are derived from the 2010 US Religious Census. The election results and a host of other county level control variables were collected by Kirkegaard (2016) and can be downloaded from OpenDataSoft.com. The following Table A.3 lists all control variables that were included in the panel regression in interaction with the indicator variable for Trump's presidential run. To analyze the correlation between Twitter Usage and the consumption of other news sources, we additionally obtain the data of the Media Consumption Survey conducted by PEW Research (available here.

Table A.3: Control Variables Used in Regression

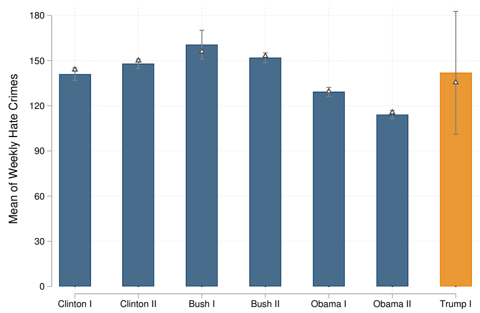
Voting Controls		**Ethnicity Controls**		**Demographic Controls**		**Economic Controls**		**Crime Controls**	
Variable	Corr.	Variable	Corr.	Variable	Corr.	Variable	Corr.	Variable	Corr.
Republican vote dummy 2008	-0.240	Share White	-0.123	Population 2016	0.260	Unemployment	-0.055	Violent Crime	0.177
Republican vote dummy 2012	-0.242	Share Black	0.079	ln(1+Population 2016)	0.273	Gini Coefficient	0.227	Homicide Rate	0.059
Republican vote dummy 2016	-0.312	Share Hispanic	0.065	Share Male	-0.065	Uninsured	-0.043		
		Share Asian	0.276	Share Age 0-4	0.265	Median Earnings 2010	0.052		
		Share American Indian	-0.047	Share Age 5-9	-0.115	Share Management	0.298		
		Share Muslims	0.156	Share Age 10-14	-0.200	Share Service	0.078		
				Share Age 15-19	0.280	Share Sales and Office	0.144		
				Share Age 20-24	0.476	Share Farming	-0.128		
				Share Age 25-29	0.306	Share Construction	-0.222		
				Share Age 30-34	0.229	Share Productions	-0.263		
				Share Age 35-39	0.106	Child Poverty	-0.091		
				Share Age 40-44	0.012	Adults over 65 in Poverty	-0.121		
				Share Age 45-49	-0.062	Share Poverty	0.012		
				Share Age 50-54	-0.215				
				Share Age 55-59	-0.254				
				Share Age 60-64	-0.199				
				Share Age 65-69	-0.201				
				Share Age 70-74	-0.230				
				Share Age 75-79	-0.252				
				Share Age 80-84	-0.222				
				Share Age 85-over	-0.134				

Notes: This table lists the controls included in the main panel regressions in each category. The second column list the correlation of the control variable with the high Twitter usage dummy ($I[90th\ Pct.\ Twitter\ Usage]$) used in the main regressions.

A.2. Supplementary Materials 2: Trends in Hate Crime by President

In this section, we provide some additional evidence on time series trends in hate crimes across US presidencies since 1990. Figure A.4 plots the average number of hate crimes per week (with the median as triangles) for each presidency. Panel a show the full period for each of the previous presidencies, while Panel b plots the weekly number of Hate Crimes until the end of the first year for each president. As mentioned in the main text, Trump does not appear to be an outlier, independently of the time interval we consider.

(a) All Hate Crimes



(b) All Hate Crimes (Equal Periods)

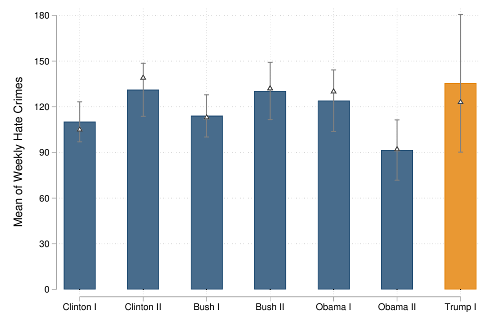


Figure A.4: Weekly Hate Crimes in the United States since 1990, by President: This figure plots the average weekly number of hate crimes, by president. Panel (a) shows the full period of the presidency. Panel (b) shows the bar graphs for the identical end-of-year periods after the election. The bars indicate the 95% confidence intervals. The triangle marks the median.

A potential issue with our comparison might be that we consider all hate crimes jointly, which could potentially hide underlying heterogeneous hate crime trends across groups. We thus reproduce the bar graphs using the other main categories of hate crimes in the FBI data (see Figure A.5). Overall, the results yield a qualitatively similar conclusion. Again Trump does not appear to be an outlier for any of the main categories.

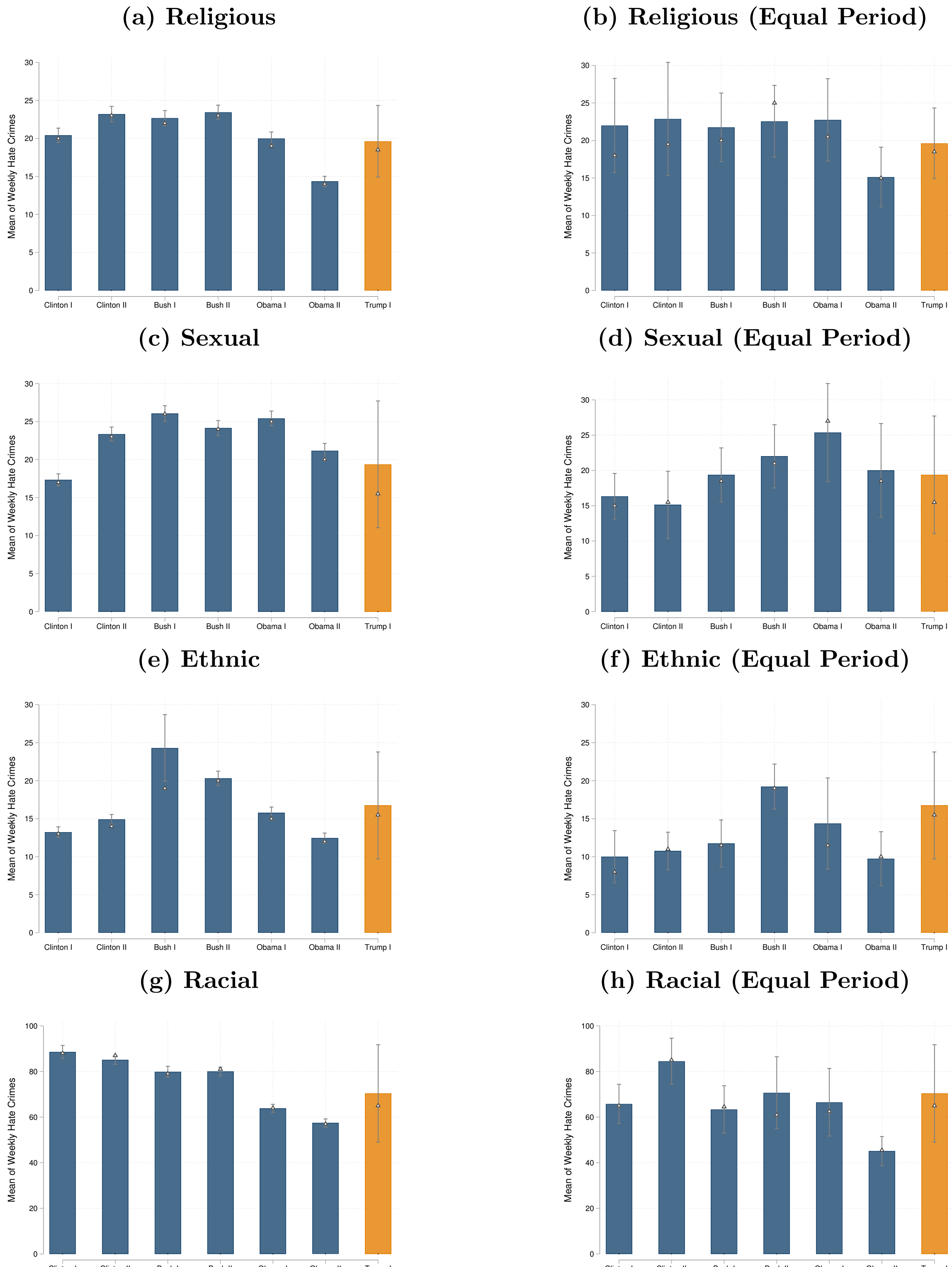


Figure A.5: Weekly Hate Crimes in the United States since 1990, by President and Motivating Bias: This figure plots the average weekly number of hate crimes, by president and type of hate crime (as defined by the FBI). Panel (a) shows the full period of the presidency. Panel (b) shows the bar graphs for the identical end-of-year periods after the election. The bars indicate the 95% confidence intervals. The triangle marks the median.

A.3. Supplementary Materials 3: Additional Time Series Evidence

In the main body of the paper, we provide the results of simple time series regressions to describe the relationship between Donald Trump's Twitter activity and hate crimes in the United States. To provide some background for these estimations, Table A.4 plots descriptive statistics on the number of hate crimes and Trump tweets for the 399 weeks since Trump's first tweet and the end of the available data in December 2016.

The data reveal a few interesting patterns. To start, the average number of weekly hate crimes has not significantly changed with the start of Trump's campaign in June 2015 (compared to the period since Trump's initial tweet in 2009). The composition of hate crimes, however, has shifted. In line with the numbers by presidency reported in the main paper, the number of anti-Muslim crimes has approximately doubled in the post-campaign period (from a median of 3 to 6 crimes per week), met mainly by a decline in ethnicity-based crimes (from a median of 21 to 15) as well as smaller decreases in sex-based and non-Muslim religious crimes. Interestingly, we also find changes in Trump's Twitter activity. The total number of his tweets increased from a median of 44 to around 94 tweets. We also see changes in content: Trump considerably intensified his messages targeting Muslims and Hispanics, and switched attention from Barrack Obama to Hillary Clinton.

Table A.4: Descriptive Time Series Statistics

	Since Trump's first tweet	Campaign start (June 16, 2015) Before	After	t-test Before = After?
Hate Crimes				
Total	118.669 [119]	118.997 [121]	117.363 [116]	0.567
Anti-Muslim	3.514 [3]	2.909 [3]	5.925 [6]	-6.7***
Anti-Ethnic Crime	20.652 [20]	21.605 [21]	16.850 [15]	5.515***
Anti-Race Crime	56.311 [56]	56.539 [56]	55.4 [54]	0.708
Anti-Sexual Crime	22.867 [22]	23.282 [23]	21.212 [20]	2.432**
Anti-Religious Crime	19.444 [18]	19.997 [19]	17.238 [16]	3.558***
Trump Tweets				
Total	75.669 [67]	69.583 [44]	99.938 [94.5]	-4.588***
Muslim	0.614 [0]	0.386 [0]	1.525 [1]	-4.152***
Hispanic	1.328 [0]	0.596 [0]	4.25 [3]	-5.478***
Obama	6.16 [3]	6.994 [4]	2.837 [3]	6.937***
Hillary	1.985 [0]	0.273 [0]	8.813 [5]	-7.522***
Loser	0.609 [0]	0.668 [0]	0.375 [0]	2.518**
Observations	399	319	80	

Notes: This table plots descriptive statistics for the weeks since Donald Trump's first tweet on May 4th, 2009 until the end of the available hate crime data on December, 31st 2016. Columns 2 through 4 present the arithmetic mean and [median] (in square brackets) for each variable shown in the left column. The right column shows a t-test for the equality of means for the variables before and after Trump's campaign start (allowing for unequal variance in the two sub-samples). *** and ** indicate statistical significance at the 0.01 and 0.05 level, respectively.

We exploit weekly variation in the content of his tweets and link them to US-wide hate crimes against Muslims in a simple time-series framework of the following type:

$$Muslim\,Hate\,Crime_{w+h} = \alpha + \beta Muslim\,Trump\,Tweets_w + \gamma Muslim\,Hate\,Crime_{w-1} + \epsilon_w. \tag{A.1}$$

The coefficient β thus captures the extent to which hate crimes in week $w+h$ react to social media salience, as proxied by Trump's tweets about Muslim-related topics. We consider the forecasting horizons 0 through 7.[5] The inclusion of a lagged dependent variable ($MuslimHateCrime_{w-1}$) is motivated by a thorough analysis of the auto-regressive properties of anti-Muslim hate crimes both before and after Trump's campaign start; model selection using the Bayesian Information Criterion (BIC) suggests that an AR(1) model is most appropriate. However, as we will show below, the results in fact vary little across different model specifications. In all regressions, we base inference about statistical significance on Newey-West standard errors to allow for heteroscedasticity and autocorrelation up to a lag of 4 weeks for the pre-campaign and 6 weeks for the post-campaign period. These values are chosen using the suggested optimal lag length of $4 \times (T/100)^{2/9}$ in Andrews and Monahan (1992). In practice, these standard errors are almost equivalent to those obtained using Huber-White standard errors or different approaches for allowing for arbitrary correlation within a given month or year.

In principle, it might be preferable to use *re-tweets* rather than Trump's own posts to capture Twitter sentiment, because they capture how influential an individual tweet is. However, we do not observe when exactly users re-tweet Trump's messages and some may do so with a substantial time lag; as such, re-tweets are more of an *ex-post* measure of Trump's influence. However, we consider re-tweets in a robustness checks below.

Appendix A.3. plots the forecasting ability of Trump's tweets on Muslim-related topic at different horizons h from table 1 in the main paper, as well as the associated 90% confidence intervals. It also visualizes the effect for the period before and after his campaign start. As we can see, there is a striking quantitative (and statistically significant) difference between the two periods. Table A.5 plots the estimates for the contemporaneous correlation of Trump tweets about Muslims and other types of hate crimes, which are shown in the main paper in figure 4. This amounts to a simple counterfactual test: If hate crimes were driven by general anti-minority factors correlated with Trump's twitter activity, we would expect them to affect violence against multiple groups. For example, Trump tweets targeting Muslims

[5] As shown in the tables in the main paper and the graphs below, the predictive ability of Trump's tweets vanishes after 3 or 4 weeks into the future. Horizons of longer than 7 weeks are always statistically insignificant (unreported).

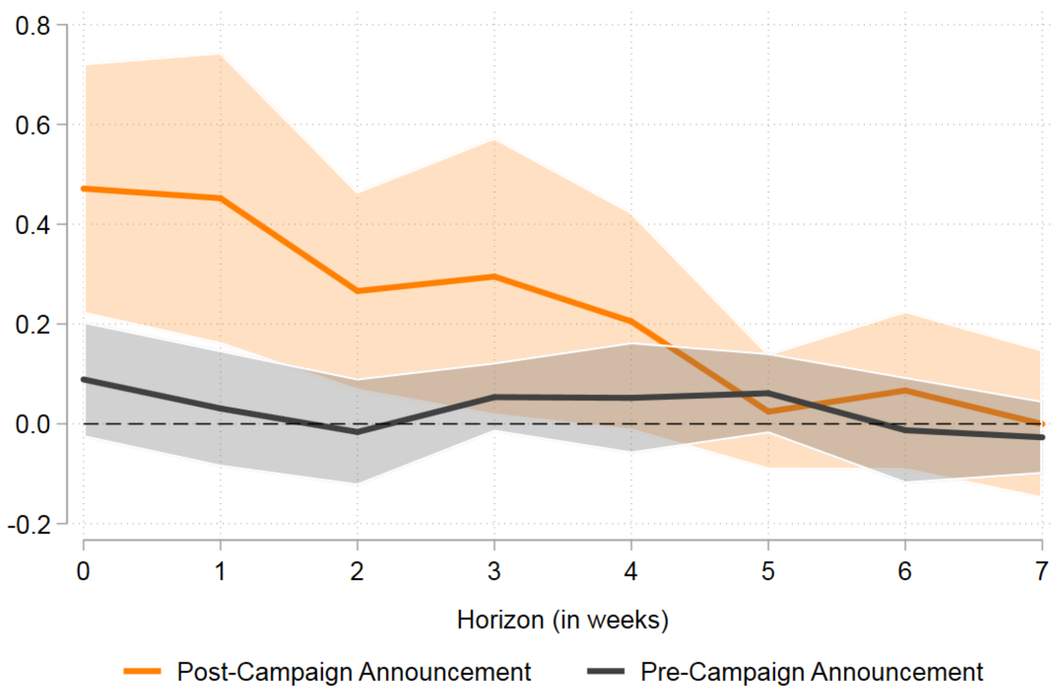


Figure A.6: Dynamic correlation of Islam-related Trump tweets and anti-Muslim hate crimes: This figure plots the standardized coefficient estimates $\hat{beta}$ of running A.1 for horizons h 0 through 7 separately for the pre-campaign and post-campaign period.

might reflect waves of violent anger towards minorities more broadly. However, we find that the explanatory power of Trump's tweets is specific to anti-Muslim hate crimes.

In Table A.6, we assess the robustness of our baseline findings. For convenience, column (1) plots the specification reported in the main body of the paper. Column (2) drops the lagged dependent variable from the specification, which in fact slightly increases the point estimate. Column (3) adds a full set of year and month-of-year dummies, and column (4) the even more stringent week-of-year dummies. Both specifications yield very similar coefficients to the baseline result. In column (5), we add the first principal component of four time series of Google searches containing the words "muslim", "islam", "mosque", and "terror" as a proxy for general attention paid to Islam-related topics in the United States. This slightly decreases the coefficient estimate but leaves its statistical significance intact. Column (6) re-runs the models in first-differences, that is a regression of the weekly *change* in anti-Muslim hate crimes on the *change* in the number of Trump tweets. This is to accommodate the concern that both hate crimes and Trump tweets are autocorrelated. Column (7) replaces the number of tweets with the number of re-tweets on Muslim-related posts by Trump. Again, these alternative specifications do not change the initial results.

How stable is the predictive ability of Trump's twitter posts? One agnostic approach to tackle this question is to run regressions using an expanding window of observations. We begin by regressing anti-Muslim hate crimes on Trump tweets for the first 30 weeks of the post-campaign period and then subsequently add the following weeks until the end of the sample on December 31, 2016. An intuitive interpretation of this procedure is that we allow the data to speak on whether there is a turning point in the relationship between tweets and hate crimes. Appendix A.3. plots the result of this exercise, where we show the changing point estimates and corresponding 90% confidence intervals. As it turns out, the link is extraordinarily stable over time.

Overall, we find that Donald Trump's tweets targeting Muslims predict hate crimes in the United States. While the time series correlations are strong, we have to be cautious about their interpretation in isolation. In particular, it remains possible that Trump's Twitter messages are themselves an outcome of waves in weekly variation in anti-minority salience. But our findings do suggest that social media and real-life actions are intimately linked.

In Table A.7, we take this finding a step further by investigating whether Trump's tweets have a higher correlation with hate crimes in counties where many people use Twitter. More precisely, we split counties into those above and beyond the median (columns (1) and (2)), 75th percentile (columns (3) and (4)), or 90th percentile (columns (5) and (6)) in the ratio of Twitter usage to population. We then re-estimate our baseline time series regression in the post-campaign period and standardize the number of anti-Muslim hate crimes to have

Table A.5: Time Series Evidence: Trump Tweets and Unrelated Hate Crimes

Dependent Variable	**Other Hate Crimes After Campaign Announcement (t)**			
	Anti-Ethnic Crime	Anti-Race Crime	Anti-Sex Crime	Anti-Religious Crime
	(1)	(2)	(3)	(4)
Muslim Trump Tweet	-0.100	0.039	0.100	-0.006
	(0.081)	(0.060)	(0.105)	(0.123)
Observations	80	80	80	80
Muslim Tweet Partial R^2	0.010	0.001	0.010	0.000
Adj. R^2	0.134	0.311	0.237	0.164

Notes: This table presents the estimated coefficients from an Ordinary Least Squares regression of contemporaneous weekly hate crimes against the groups indicated in the top row on tweets by Donald Trump containing Muslim-related words. All variables are standardized with mean 0 and standard deviation 1 and thus directly comparable across the regressions and with the results in Table A.5. The sample includes the period after Donald Trump's campaign announcement until the end of the available hate crime data on December 31, 2016. All regressions include one lag of the dependent variable, as suggested by a lag selection test based on the Bayesian information criterion (BIC). The Partial R^2 is calculated as the difference of the regression model with and without the Muslim Trump Tweet variable. Newey-West standard errors are reported in parentheses, allowing for 4 lags, in line with the lag selection procedure in Andrews and Monahan (1992). ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.1 level, respectively.

Table A.6: Robustness: Trump Tweets and Anti-Muslim Hate Crimes

	Baseline Specification (1)	No Lagged Dependent Variable (2)	Add Year And Month Dummies (3)	Add Year And Week-of-Year Dummies (4)	Baseline With Google Search Control (5)	Use First-Differences (6)	Use Trump Retweets (7)
Panel A: Before Campaign Announcement							
Muslim Trump Tweet	0.089	0.086	0.084	0.082	0.004	0.100	0.068
	(0.071)	(0.074)	(0.068)	(0.067)	(0.065)	(0.071)	(0.042)
Observations	318	319	305	305	139	318	318
Adj. R^2	0.029	0.007	0.179	0.299	0.191	0.010	0.026
Panel B: After Campaign Announcement							
Muslim Trump Tweet	0.471***	0.499***	0.405***	0.433**	0.336**	0.375**	0.365***
	(0.153)	(0.177)	(0.145)	(0.180)	(0.161)	(0.172)	(0.133)
Observations	80	80	80	56	80	80	80
Adj. R^2	0.433	0.249	0.526	0.757	0.459	0.140	0.345

Notes: This table presents the estimated coefficients from an Ordinary Least Squares regression of weekly hate crimes at the horizon indicated in the top row on tweets by Donald Trump containing Muslim-related words. All variables are standardized with mean 0 and standard deviation 1 and thus directly comparable across the regressions. In Panel A, we only include the period before Donald Trump announced his presidential campaign. Panel B includes the period after his announcement until the end of the available hate crime data on December 31, 2016. (2) drops the lagged dependent variable. (3) adds year and month-of-year dummies. (4) adds year and week-of-year dummies. (5) adds as a control variable the principal component of weekly Google Searches for the words "muslim", "islam", "mosque", and "terror". (6) re-runs the specification in column (2) in first-differences. (7) uses re-tweet numbers instead of re-tweets. Newey-West standard errors are reported in parentheses, allowing for 4 and 6 lags in Panel A and B, respectively, in line with the lag selection procedure in Andrews and Monahan (1992). ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.1 level, respectively.

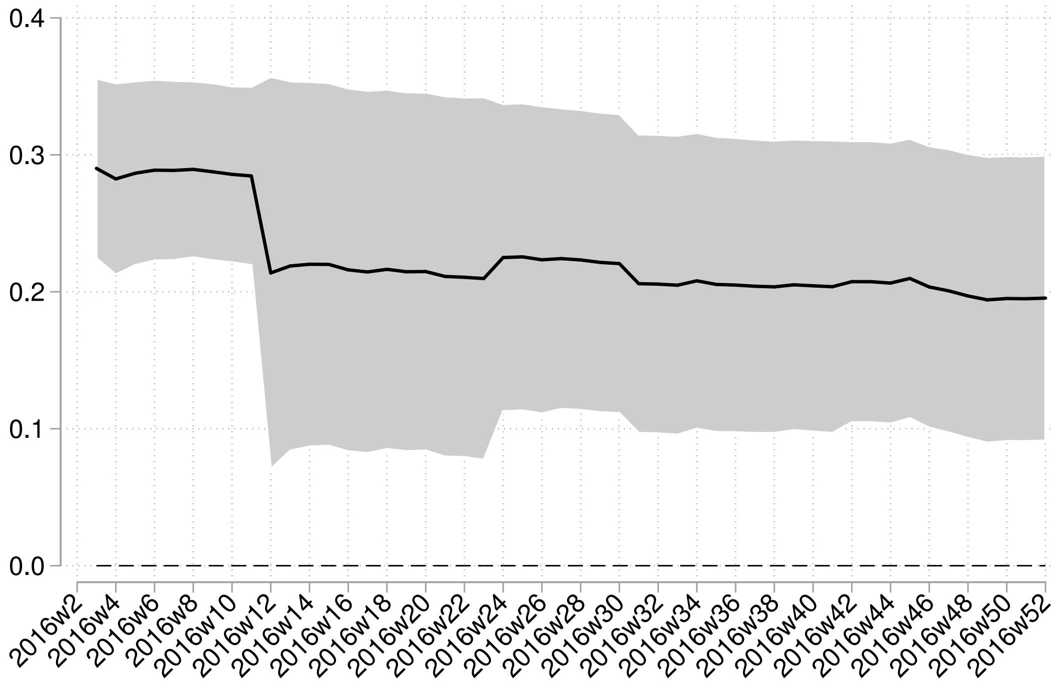


Figure A.7: Expanding window regressions of anti-Muslim hate crimes on contemporaneous Trump tweets on Muslim-related topics

mean zero and standard deviation of 1; as such, we abstract from any differences in the average number and volatility of hate crimes in areas depending on Twitter usage.

The results suggest that Trump's tweets about Islam-related topics show considerably higher correlations in counties with high Twitter usage throughout. For example, the correlation has a coefficient of 0.168 in counties above the median (column 1) but is basically zero below the median (column 2). This is also reflected in the explanatory power for anti-Muslim hate crimes, as measured by their "partial R^2" (i.e. the explanatory power over and above the lagged dependent variable we continue to include in the model). When we split counties based on the median, Trump's tweets basically only have explanatory power where many people use Twitter. The partial R^2 is around 3.6 times and 40% higher in counties with high Twitter usage for the top quartile and top decile sample split, respectively. These results are intuitive, given that we increasingly restrict the sample to fewer and fewer counties (which yields fewer variation in hate crimes). While they can only be taken as suggestive, they also highlight that the local presence of Twitter increases the transmission of social media to real-life action.

Table A.7: Time Series Evidence by Twitter Usage: Trump Tweets and Hate Crimes

	Twitter Usage/Pop. Measure					
	Median		75th Percentile		90th Percentile	
	High	Low	High	Low	High	Low
Muslim Trump Tweet	0.168***	-0.009	0.159***	0.083	0.154***	0.131***
	(0.050)	(0.051)	(0.048)	(0.056)	(0.057)	(0.044)
Observations	80	80	80	80	80	80
Muslim Tweet Partial R^2	0.161	0.000	0.144	0.039	0.135	0.099
Adj. R^2	0.282	0.064	0.274	0.040	0.250	0.110

Notes: This table presents the estimated coefficients from an Ordinary Least Squares regression of contemporaneous weekly hate crimes against the groups indicated in the top row on tweets by Donald Trump containing Muslim-related words, split by counties with "many" and "few" Twitter users (as indicated in the top row). All variables are standardized with mean 0 and standard deviation 1 within the group of many and few Twitter users and thus directly comparable across the regressions. The sample includes the period after Donald Trump's campaign announcement until the end of the available hate crime data on December 31, 2016. All regressions include one lag of the dependent variable, as suggested by a lag selection test based on the Bayesian information criterion (BIC). The Partial R^2 is calculated as the difference of the regression model with and without the Muslim Trump Tweet variable. Newey-West standard errors are reported in parentheses, allowing for 4 lags, in line with the lag selection procedure in Andrews and Monahan (1992). ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.1 level, respectively.

A.4. Supplementary Materials 4: Additional Panel Evidence

In this section, we provide additional panel evidence in line with the results presented in the main body of the paper. We start by documenting that the finding of a disproportionate increase in anti-Muslim hate crimes in areas with many Twitter users is not driven by the choice of comparing counties in the top 10 and bottom 90 percent of the ratio of Twitter users to population. To this end, Figure A.8 and instead split the ratio into the 75-25 percentiles. Figure A.9 shows the the split at the 90th percentile for the unscaled raw data without any fixed effects. This yields a highly similar visual pattern as our main results.

Next, we conduct a more formal diagnostic test of whether hate crimes prior to Trump's presidential campaign announcement followed similar trajectories in areas with many and few Twitter users. This corresponds to the parallel trends assumption required for causal inference in difference-in-differences regressions (Angrist and Pischke, 2008). Note, however, that we do not interpret our findings as evidence of a causal effect of social media on hate crime; rather, we believe they simply document a striking pattern in the data, which we believe future research should address more rigorously to identify causal effects (also see Müller and Schwarz, 2018). In particular, we include leads of the "Pres Run" dummy that tags the period after Donald Trump's presidential run in our baseline panel regressions. For example,

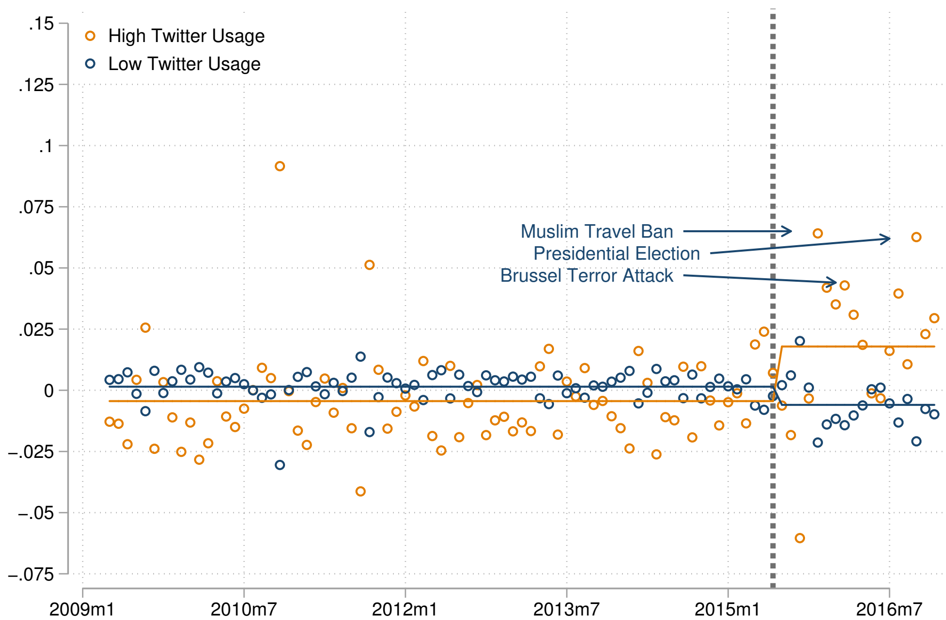


Figure A.8: Monthly Anti-Muslim Hate Crimes per capita by Twitter Usage: This figure plots the monthly number of anti-Muslim hate crimes for counties above and below the 75th percentile of Twitter usage; data are adjusted for county and week fixed effects. The vertical line marks the beginning of Trump's presidential campaign.

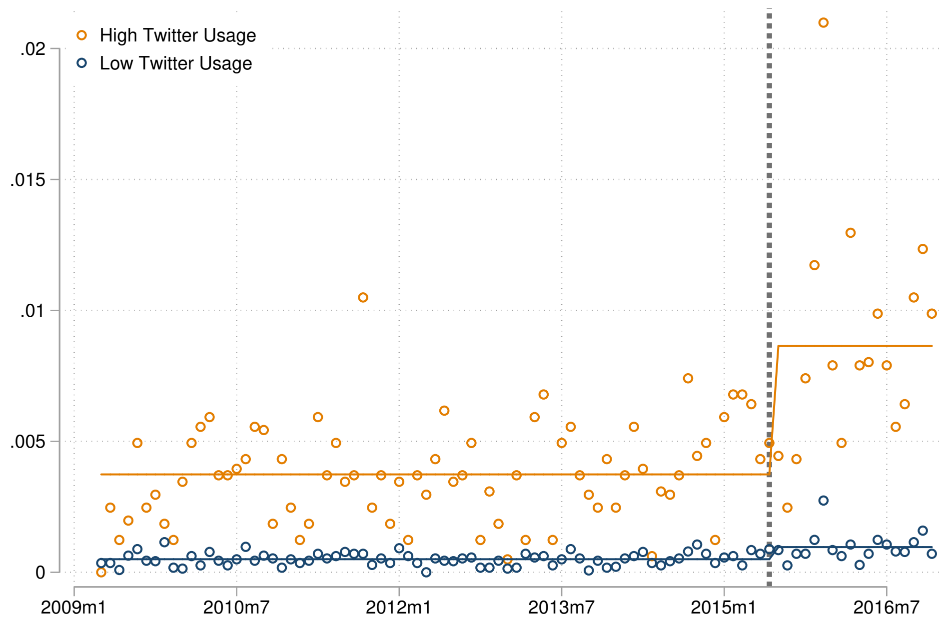


Figure A.9: Monthly Anti-Muslim Hate Crimes by Twitter Usage: This figure plots the monthly number of anti-Muslim hate crimes for counties above and below the 90th percentile of Twitter usage. The vertical line marks the beginning of Trump's presidential campaign.

Pres Run Lead 1 tags the period 1 year in advance of Trump's presidential run. Statistically significant estimates of the parameters of these terms would indicate that anti-Muslim hate crimes already followed different trends prior to the campaign start. However, the results we plot in columns (1) through (3) in Table A.8 suggest that such concerns are not borne out in the data.

Next, we show a test in the spirit of Figure 3a and Figure A.9 to rule out that our results are driven by the functional form of the regression. For these, we replace our Twitter usage measure based with the *Twitter usage per capita*, indicator variables for the 50th, 75th, 95th percentile and $Ln(1 + TwitterUsage)$. For all of the alternative Twitter usage measures we again find significant positive effects for the period after Trump's presidential campaign.

In Table A.9, we replace the dependent variable with hate crimes based on other motivations: ethnicity, race, sexual orientation, and religion. In line with our time series results, we do not find a disproportional change in hate crimes based on racial, sexual, or non-Muslim religious motivations in areas with higher Twitter usage around Trump's presidential campaign announcement. This supports our previous results that the change in hate crime under Trump was predominantly related to Muslims – the minority that arguably most unapologetically in his verbal cross-hair. However, we do find an increase in anti-Ethnic

Table A.8: Additional Panel Results

	Testing for Pretends			Alternative Twitter Usage Measures				
Dependent Variable	Anti-Muslim Hate Crime			Twitter Usage per Capita	50th Percentile	75th Percentile	95th Percentile	Ln(1+Twitter Usage)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$Pres\ Run \times$	0.040**	0.039**	0.037**	0.010**	0.023***	0.033***	0.030*	0.0004***
$I[90th\ Pct.\ Twitter\ Usage]$	(0.017)	(0.017)	(0.018)	(0.004)	(0.008)	(0.012)	(0.017)	(0.0001)
$Pres\ Run\ Lag\ 1 \times$	0.011							
$I[90th\ Pct.\ Twitter\ Usage]$	(0.008)							
$Pres\ Run\ Lag\ 2 \times$		0.003						
$I[90th\ Pct.\ Twitter\ Usage]$		(0.006)						
$Pres\ Run\ Lag\ 3 \times$			-0.003					
$I[90th\ Pct.\ Twitter\ Usage]$			(0.007)					
State-Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,253,259	1,253,259	1,253,259	1,253,259	1,253,259	1,253,259	1,253,259	1,253,259
Number of counties	3,141	3,141	3,141	3,141	3,141	3,141	3,141	3,141
R-squared	0.022	0.022	0.022	0.022	0.022	0.022	0.022	0.0678

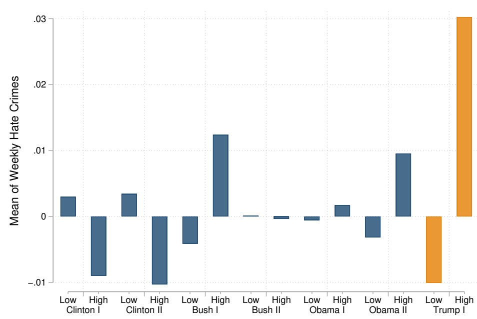
Notes: This table presents the estimated coefficients from a regression of hate crimes on Twitter usage described in the text. The dependent variable is the number of hate crimes committed against Muslims per capita. The dependent variable is standardized with mean 0 and standard deviation 1. *Pres Run* is an indicator variable for the time period from the beginning of Donald Trump's presidential run. *Pres Run Lead t* is an indicator variable for the t years before Donald Trump's presidential run. $I[90th\ Pct.\ Twitter\ Usage]$ is an indicator variable that is 1 for counties above the 90th percentile of Twitter usage per capita. All regression include county and state-week fixed effects. The regression using Ln(1+Twitter Usage) as a measure of Twitter usage additional controls for the population and the logarithm of the population. Robust standard errors in all specifications are clustered by county. ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.1 level, respectively.

hate crimes (albeit statistically insignificant): the estimated coefficient of 0.033 is considerably larger than that for anti-racial or other anti-religious, and similar to the coefficient of 0.038 for anti-Muslim hate crimes. As we will document below, a careful analysis of the data reveals that that there is indeed a similar but somewhat weaker link between Trump's tweets about Hispanics and hate crimes on Ethnic grounds, which usually target Hispanics (see Supplementary Material 6).

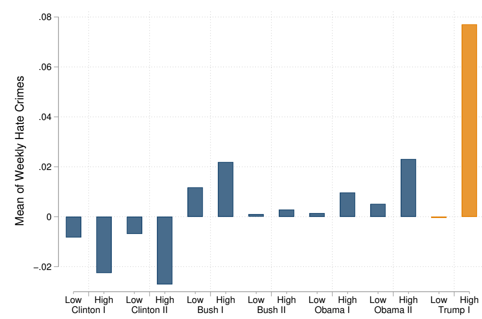
A.5. Supplementary Materials 5: Comparison with other Presidencies

In this section, we provide some robustness exercises on the number of anti-Muslim hate crimes in areas with many and few Twitter users by presidency. In particular, Figure A.10 shows four alternative ways of displaying these data. In panel (a), we split counties on the 75th percentile instead of 90th percentile, which turns out to make only a small difference for the results. In panel (b), we then limit the data adjustment to county fixed effects, which de-means the values with their county-specific average. Again, Trump clearly stands out in comparison.

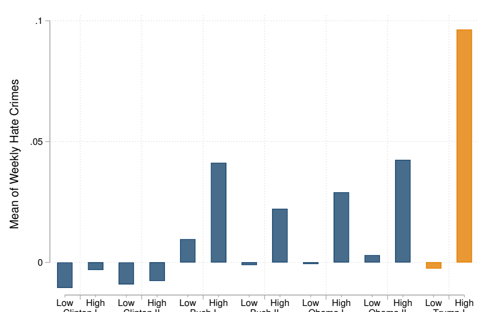
(a) 75th Percentile



(b) Only County FE



(c) No FE



(d) Not Standardized

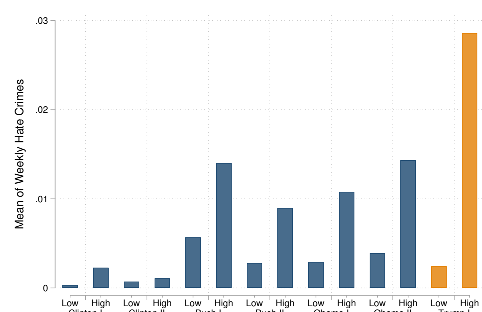


Figure A.10: Number of anti-Muslim Hate Crimes per Capita, by President and Twitter Usage: The figures show the average weekly number of anti-Muslim hate crimes per capita split by Twitter usage and presidency. Panel (a) splits counties at the 75th percentile of Twitter usage. Panel (b) only conditions on county fixed effects, Panel (c) shows the figure without fixed effects and panel (d) shows the unstandardized figure without fixed effects.

Finally, in panel (c), we present (standardized) values of the raw data without any adjustment. Again, the difference in anti-Muslim hate crimes is clearly most concentrated in counties with many Twitter users under Trump compared to all other presidencies since 1990, even without abstract from time or county-specific determinants. In panel (d), we repeat the figure in panel (c) with unstandardized form, i.e. the raw number of anti-Muslim hate crimes per million inhabitants.

Table A.9: Shifts in Other Hate Crimes in Panel Regressions

Dependent Variable	Anti-Ethnic (1)	Anti-Ethnic (2)	Anti-Racial (4)	Anti-Racial (4)	Anti-Sexual (5)	Anti-Sexual (6)	Anti-Religious (7)	Anti-Religious (8)
$Pres\ Run \times$	0.033		-0.019**		-0.005		-0.000	
$I[90th\ Pct.\ Twitter\ Usage]$	(0.033)		(0.009)		(0.008)		(0.004)	
$Pres\ Run \times$		0.019		0.004		-0.001		0.004
$I[75th\ Pct.\ Twitter\ Usage]$		(0.015)		(0.007)		(0.005)		(0.004)
State-Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,253,259	1,253,259	1,253,259	1,253,259	1,253,259	1,253,259	1,253,259	1,253,259
Number of counties	3,141	3,141	3,141	3,141	3,141	3,141	3,141	3,141
R-squared	0.020	0.020	0.027	0.027	0.021	0.021	0.028	0.028

Notes: This table presents the estimated coefficients from a regression of hate crimes on Twitter usage described in the text. The dependent variables are the number of hate crimes committed against the given minority per capita. The dependent variables are standardized with mean 0 and standard deviation 1. $Pres\ Run$ is an indicator variable for the time period from the beginning of Donald Trump's presidential run. $I[90th\ Pct.\ Twitter\ Usage]$ is an indicator variable that is 1 for counties above the 90th percentile of Twitter usage per capita. Similarly, $I[75th\ Pct.\ Twitter\ Usage]$ is an indicator variable for the 75th percentile. All regression include county and state-week fixed effects. Robust standard errors in all specifications are clustered by county. ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.1 level, respectively.

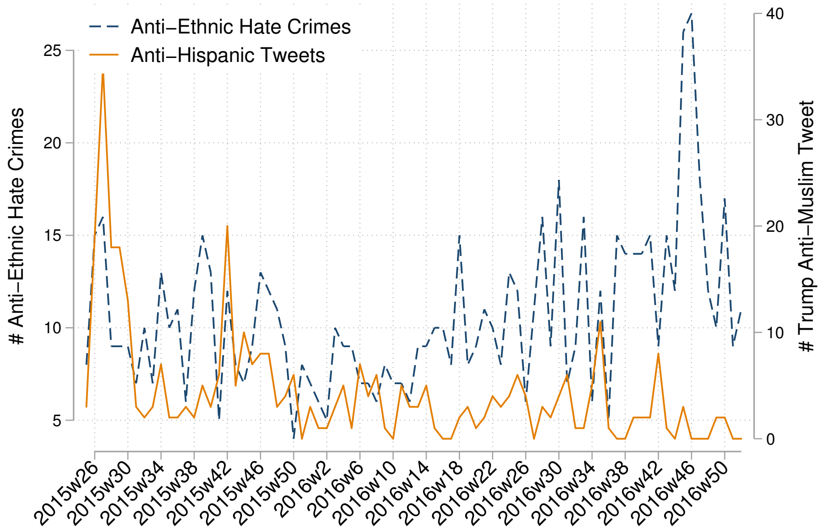


Figure A.11: Trump Tweets about Hispanics and Ethnicity-Motivated Hate Crime: This figure shows the number of tweets by Donald Trump about Hispanics and the number of anti-Ethnic hate crimes in a given week.

A.6. Supplementary Materials 6: Additional Evidence – Hispanic Tweets and Anti-Ethnic Crimes

In the main body of the paper, we have provided evidence that Trump's tweets about Muslims are highly correlated with hate crimes against that particular group. Here, we provide some additional evidence based on Trump's twitter posts aimed at Hispanics. The correlations we document here are somewhat weaker than those for Muslims, which appears to be driven mainly by a change in Trump's tweeting behavior in the weeks in the run-up to the 2016 election. The evidence that anti-Ethnic hate crimes have become more concentrated in areas with many Twitter users in the period after Trump's campaign announcement is also somewhat weaker. Nevertheless, we believe our results paint a complementary picture to the link between Donald Trump's Twitter activity and the time series of hate crimes in the United States.

Figure A.11 visualizes the time series pattern after Trump's campaign announcement that we will investigate more formally in a time series regression framework (as in eq. (A.1)). In particular, we now look at the number of ethnicity-based hate crimes.[6] For Trump's tweets, we focus on posts that include the words "latino", "hispanic", "immigra", "illegal alien",

[6]The FBI hate crime data further differentiate between "Hispanic" and "Other" ethnicity-based crimes. In unreported regressions, we also replicated the results using only anti-Hispanic hate crimes, which yields similar but statistically weaker results. We suspect that the more comprehensive ethnicity-based classification is more accurately measured, which may explain the more precisely estimated results.

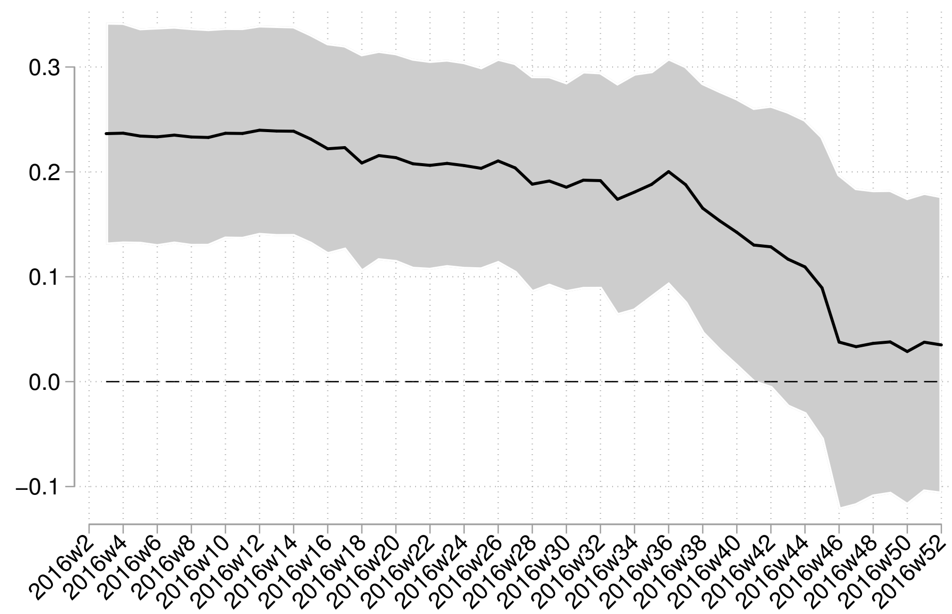


Figure A.12: Expanding window regressions of anti-Ethnic hate crimes on contemporaneous Trump tweets on Hispanic-related topics

"border", and "wall".[7] The figure suggests a systematic correlation between Trump's tweets aimed at Hispanics and ethnicity-based hate crimes, mainly for the first year until mid-2016.

We put this correlation to a statistical test in table A.10. As before, we standardize all variables by their sub-samples to have a mean of 0 and standard deviation of 1; this makes the estimated coefficients directly comparable. In Panel A, we begin by considering the baseline regression in levels, i.e. the raw numbers of attacks and Twitter posts. Column (1) shows that before the campaign start, Trump's tweets were uncorrelated with anti-Hispanic hate crimes. The same is true for the specification in Panel B, where we take the change of both hate crimes and tweets as first-differences. In column (2), we consider the 54 weeks between Trump's campaign start and the last weeks prior to his election; we will discuss the reason for this cut-off further down. In line with our main results on Muslims, the coefficient of 0.185 is highly statistically significant, suggesting a striking time series correlation between Trump's tweets about Hispanics and hate crimes. The coefficient for the estimation in first-differences in Panel B is almost equivalent.

Next, we expand the period to the entire post-campaign start sample in column (3). The regression in levels now yields a much smaller (but positive) coefficient, which is not statistically significant at conventional levels. In first-differences, however, the coefficient remains unchanged and now gains statistical power (significant at the 5% level).

[7]As in the case of the Muslim tweets, the exact word choice makes little difference to our results.

Table A.10: Time Series Evidence: Trump Tweets and Ethnicity-Based Hate Crimes

	Before Campaign Start (1)	After Campaign Start (Pre-Election) (2)	After Campaign Start (Full) (3)
Panel A: Levels			
Hispanic Trump Tweet	-0.001	0.204***	0.035
	(0.049)	(0.061)	(0.086)
Observations	318	55	80
Adj. R^2	0.063	0.141	0.125
Panel B: First-Differences			
Hispanic Trump Tweet	0.033	0.208***	0.183**
	(0.044)	(0.069)	(0.075)
Observations	317	55	80
Adj. R^2	0.266	0.287	0.249

Notes: This table presents the estimated coefficients from an Ordinary Least Squares regression of weekly ethnicity-based hate crimes on tweets by Donald Trump containing Hispanic-related words. All variables are standardized with mean 0 and standard deviation 1 and thus directly comparable across the regressions. In Panel A, we use all variables in levels; in Panel B, in first-differences. Column (1) only includes the period before Donald Trump announced his presidential campaign (the week of June 16, 2015). Column (2) includes the period after his announcement until week 26 of 2016 (excluding the weeks immediately preceding his election as president). Column (3) includes the entire period until the end of the available hate crime data on December 31, 2016. All regressions include one lag of the dependent variable, as suggested by a lag selection test based on the Bayesian information criterion (BIC). Newey-West standard errors are reported in parentheses, allowing for 6 lags before and 4 lags after the campaign start, in line with the lag selection procedure in Andrews and Monahan (1992). ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.1 level, respectively.

What may explain this decaying effect? One agnostic approach to tackle this question is to run regressions using an expanding window of observations. We begin by regressing anti-Ethnic hate crimes on Trump tweets for the first 30 weeks of the post-campaign period and then subsequently add the following weeks until the end of the sample on December 31, 2016. An intuitive interpretation of this procedure is that we allow the data to speak on a turning point in the relationship between tweets and hate crimes. Figure A.12 plots the result of this exercise, where we show the changing point estimates and corresponding 90% confidence intervals. The emerging picture is that Trump's Hispanic tweets and contemporaneous ethnicity-based hate crimes are significantly correlated at least at the 10% level until week 39 of 2016, with a relatively stable (standardized) coefficient of around 0.2. What explains the subsequent drop? A look back at Figure A.11 reveals that in the weeks leading up to his election as president, Trump's Twitter activity did not coincide with major bursts of anti-Ethnic hate crimes. A thorough reading of his Twitter feed suggests that this mainly reflects his changed preferences for promoting his extensive campaign touring; this is also consistent with contemporaneous media reports that asked whether Trump had become "more presidential" in preparation for a potential election win.[8]

Table A.11: Trump's Hispanic-Related Tweets and Other Hate Crimes

	Other Hate Crimes After Campaign Announcement (t)			
	Anti-Muslim Crime	Anti-Race Crime	Anti-Sex Crime	Anti-Religious Crime
	(1)	(2)	(3)	(4)
Hispanic Trump Tweet	-0.086	-0.033	0.136	-0.090
	(0.063)	(0.054)	(0.116)	(0.072)
Observations	80	80	80	80
Adj. R^2	0.219	0.311	0.243	0.172

Notes: This table presents the estimated coefficients from an Ordinary Least Squares regression of contemporaneous weekly hate crimes against the groups indicated in the top row on tweets by Donald Trump containing Hispanic-related words. All variables are standardized with mean 0 and standard deviation 1 and thus directly comparable across the regressions and with the results in Table A.10. The sample includes the period after Donald Trump's campaign announcement until the end of the available hate crime data on December 31, 2016. All regressions include one lag of the dependent variable, as suggested by a lag selection test based on the Bayesian information criterion (BIC). Newey-West standard errors are reported in parentheses, allowing for 4 lags, in line with the lag selection procedure in Andrews and Monahan (1992). ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.1 level, respectively.

Analogous to our main results on Muslims, we conduct counterfactual exercises in Table A.11. In particular, we regress the other types of hate crimes (anti-muslim, anti-race, anti-sex, anti-religious) on Hispanic tweets in the period after the beginning of Donald Trump's presidential campaign. We do not find any significant correlation; in unreported

[8] See a late-April article on Reuters for an early example.

results, we have confirmed that the same holds true for a specification in first-differences instead of levels (as in Panel B of Table A.10).

In Table A.12, we also replicate the panel regressions from the main text for anti-Ethnic hate crimes. In particular, we test whether Ethnicity-based crimes have increased disproportionally in counties with many Twitter users. As we have seen in Table A.9 above, this increase is quantitatively similar to those for Muslims but not statistically significant when splitting Twitter users based on the 90th percentile of the user to population ratio. Here, we instead define counties above the median in the Twitter user/population ratio as having "many Twitter users". As it turns out, the increase becomes statistically significant with this definition, even after controlling a host of observable characteristics on the county-level.

Table A.12: Changes in Anti-Ethnic Hate Crime with Trump's Presidential Run

Dependent Variable	Anti-Hispanic Hate Crime					
	(1)	(2)	(4)	(5)	(6)	(7)
$Pres\ Run \times$	0.015*	0.013*	0.015*	0.030*	0.019**	0.002
$I[50th\ Pct.\ Twitter\ Usage]$	(0.008)	(0.007)	(0.009)	(0.017)	(0.009)	(0.005)
Voting Controls		Yes				
Ethnicity Controls			Yes			
Demographic Controls				Yes		
Economic Controls					Yes	
Crime Controls						Yes
State-Week FE	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,253,259	1,241,289	1,253,259	1,253,259	1,250,865	483,189
Number of counties	3,141	3,111	3,141	3,141	3,135	1,211
R-squared	0.020	0.020	0.020	0.020	0.020	0.068

Notes: This table presents the estimated coefficients from a regression of hate crimes on Twitter usage described in the text. The dependent variable is the number of hate crimes committed against Ethnic minorities per capita. The dependent variable is standardized with mean 0 and standard deviation 1. *Pres Run* is an indicator variable for the time period from the beginning of Donald Trump's presidential run. $I[50th\ Pct.\ Twitter\ Usage]$ is an indicator variable that is 1 for counties above the median of Twitter usage per capita. All regression include county and state-week fixed effects. See text for an explanation of the control variables. Robust standard errors in all specifications are clustered by county. ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.1 level, respectively.

A.7. Supplementary Materials 8: Correlations of Twitter Usage and Media Consumption

We present some evidence based on survey data from the Pew Research Center that twitter users are not disproportionately more likely to get their news from Fox News compared to other sources in Table A.13. In particular, we use Pew's Media Consumption Survey 2012, based on about 3,000 respondents, which contains the question "Do you ever use Twitter or read Twitter messages, or not?".

According to the survey, Twitter users prefer newspapers over TV and radio. Perhaps unsurprisingly, they also cite social media and Twitter as important sources for news. Within popular TV news shows, they are considerably more likely to watch MSNBC and CNN compared to Fox News, NBC, ABC, or CBS (all of which show almost equivalent correlations).

Table A.13: Media Usage Correlations

			Media Usage			
	Newspaper	TV	Radio			
Twitter Usage	0.1197	0.0508	0.0202			
			News Sources			
	News Program	News Online	News Cell Phone	News Twitter	News Social Media	
Twitter Usage	0.0762	0.148	0.1883	0.3821	0.2386	
			TV News Shows			
	CBS News	ABC News	NBC News	Fox News	MSNBC News	CNN News
Twitter Usage	0.0902	0.0818	0.0965	0.102	0.1787	0.1774

Notes: This table presents correlations of Twitter with other media and news sources. The correlations are based on the 2012 Media Consumption Survey Conducted by PEW Research. In total 3,003 people were interviewed for the survey. For the calculation of the correlations, we recode all questions as dummy variables where the answers "Regularly" and "Sometimes" were coded as 1, and "Hardly ever" and "Never" as 0. This is to make these answers comparable with "Yes/No" questions. Note that the results are similar if we recode these categories as 1 to 4.