Terrorism, Islamophobia, and Radicalization

Tamar Mitts

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2017

ABSTRACT

Terrorism, Islamophobia, and Radicalization

Tamar Mitts

Why do ordinary people become supportive of violent, extremist ideologies? Over the past several years, tens of thousands of individuals across the world have become attracted to propaganda disseminated by the Islamic State (ISIS), and many have left their home countries to join the organization. This dissertation closely examines possible explanations for pro-ISIS radicalization in Europe and the United States. I argue that anti-Muslim hostility is an important driver of pro-ISIS radicalization, leading individuals who feel isolated to become attracted to the organization's propaganda. I also contend that groups like ISIS are aware of this pattern, and thus seek to purposefully provoke hostility against potential supporters by carrying out terrorist attacks. I maintain that efforts to stop radicalization should focus on ways to reduce hostility and increase inclusion of minorities in the West. The various dissertation papers empirically examine different aspects of these arguments.

In the first paper, I examine whether anti-Muslim hostility might be driving pro-ISIS radicalization in Europe, by analyzing the online activity of thousands of ISIS sympathizers in France, Germany, Belgium, and the United Kingdom. Matching online radicalization indicators with offline data on vote share for far-right, anti-Muslim parties, I show that the intensity of anti-Muslim hostility at the local (neighborhood/municipality) level strongly correlates with support for ISIS on Twitter. In addition, I show that events that stir anti-Muslim sentiment, such as terrorist attacks and anti-Muslim protests, lead ISIS sympathizers to significantly increase pro-ISIS rhetoric, especially in areas with high far-right

support.

In the second paper, I argue that armed groups strategically use terrorism to manipulate levels of anti-Muslim hostility in Western countries. I test whether terrorism leads to greater expressions of anti-Muslim hostility using data on thirty-six terrorist attacks perpetrated by radical jihadists in the West from 2010 to 2016, examining how they shaped anti-Muslim attitudes among individuals in targeted countries. I find that individuals systematically and significantly increase posting of anti-Muslim content on social media after exposure to terrorism. The effect spikes immediately after attacks, decays over time, but remains significantly higher than pre-attack levels up to a month after the events. The results also reveal that the impact of terrorist attacks on anti-Muslim rhetoric is similar for individuals who already expressed hostility to Muslims before the attacks and those who did not. Finally, I observe that the impact of terrorist attacks on anti-Muslim hostility increases with attacks resulting in greater numbers of casualties.

In the third paper, I examine what might be done to stop online radicalization and support for ISIS in the West. I collected data on community engagement events performed in the United States by the Obama Administration, which aimed to increase trust and relationships between the Muslim population and the American government, and combined them with high-frequency, geo-located panel data on tens of thousands of individuals in America who follow Islamic State accounts on Twitter. By analyzing over 100 community engagement events in a Difference-in-Differences design, I find that community engagement activities are systematically and significantly associated with a reduction in pro-ISIS rhetoric on Twitter among individuals located in event areas. In addition, the observed negative relationship between community engagement activities and pro-ISIS rhetoric is stronger in areas that held a large number of these events.

Contents

List of Tables

List of Figures

Acknowledgments

This dissertation would not have become a reality without the help and advice of my wonderful advisers at Columbia. I would like to thank Christopher Blattman, Page Fortna, Macartan Humphreys, Jack Snyder, and Suresh Naidu for their generous advice and feedback in various stages of the project. I am grateful to Chris for his mentorship from the earliest phases of the dissertation, his commitment to give detailed feedback on numerous drafts, and his openness to my (sometimes crazy!) ideas. I thank Page for exposing me to new ways of thinking about research, and her encouragement to think globally, beyond local and familiar settings. Her advice to study the effects of terrorism in Europe was a big part for how this dissertation unfolded. I am most grateful for Macartan's thoughtful ideas and comments on various parts of the dissertation. His feedback inspired me to push the boundaries of what can be done when studying radicalization and extremist violence. I am indebted to Jack for his advice to think about the bigger picture and embrace the complexities inherent in this type of research topic. Finally, I am thankful to Suresh for inspiring me to learn about text-as-data and machine learning, and his contagious creativity and excitement about research. I am also thankful for the advice I received from Eli Berman, Jacob Shapiro, Guy Grossman, John Huber, and Joshua Tucker.

Beyond my advisers, I am deeply grateful for my friends and colleagues at Columbia, an incredible group of smart and creative individuals. I am particularly thankful to Lindsay Dolan for her thoughtful comments on my various project ideas, even from our first days at

Columbia. Her friendship, encouragement, advice and support were a big part of my ability to successfully complete this dissertation. I also thank Susanne Baltes, Jasper Cooper, Grant Gordon, Kolby Hanson, Sarah Khan, Sung Eun Kim, Egor Lazarev, Summer Lindsey, Kunaal Sharma, Camille Strauss-Kahn, Lauren Young, and the members of the Chris-Macartan advising group for their really great feedback, advice, and ideas.

But most of all, I am deeply thankful for my husband, Josh, the smartest person I know, for his incredible support, patience, and kindness. I am grateful to Josh for spending hours talking with me about my projects, his feedback on various parts of this dissertation, and his endless creative resourcefulness. Josh's wisdom and advice were paramount in helping me craft the research designs and data collection for this dissertation. And most important, I am indebted to Josh for taking care of our daughter, Emily, to give me time to work. Josh, you are everything I ever wanted and I am so thankful for the love and happiness your brought into my life.

I also thank my family for being supportive throughout the process, even though we were far away. I thank my father, Yehuda, for listening to every little project I was working on, asking questions and becoming familiar with complex academic terms. I thank my mother, Kati, for her never-ending love and kindness. To my sisters, Hadas and Shaked, I am most grateful for always being there for me, even with our crazy time-zone difference. Thank you for your support, friendship, patience, and love, and for giving me your important perspectives on the topics of my research.

To my beloved family

Chapter 1

Introduction

Over the past few years, tens of thousands of individuals across the world have become attracted to propaganda disseminated by the Islamic State (ISIS) on the Internet and social media. Many have joined the organization to become foreign fighters, leaving their home countries to wage war on ISIS's behalf in Iraq and Syria (Schmitt and Sengupta, 2015). Compared to prior conflicts, the Islamic State's recruitment of supporters stands out its scope, nature, and scale: since the beginning of the Syrian civil war, the group has recruited individuals from over 100 countries, across various regions of the world and from diverse ethnic and religious backgrounds (Barrett et al., 2015). In the West, no single profile characterizes Islamic State supporters: some were born to Muslim families, while others converted to Islam; some are immigrants, while others are native-born citizens. What can explain the vast support for ISIS in so many Western countries, in such short amount of time? What makes ordinary people attracted to such an extreme, violent ideology? And what might be done to stop this wave of radicalization?

This dissertation seeks to answer these questions in several ways. First, I argue that hostility against Muslims is an important force driving support for the Islamic State in the West. I argue that individuals who feel isolated and alienated might be attracted to the propaganda disseminated by ISIS, which can increase the likelihood that they will be radicalized and recruited to join the organization. Second, I suggest that extremist violent groups like ISIS are aware of this pattern, and thus seek to purposefully provoke hostility against potential supporters by carrying out terrorism. Third, I show that anti-Muslim hostility is strongly correlated with support for far-right parties in the West, and that pro-ISIS extremism and far-right support feed each other in a powerful vicious cycle. I suggest that efforts to stop radicalization should focus on ways to reduce hostility and increase inclusion of minorities in the West.

To test these theoretical arguments, I collected original data on the online behavior of over a million Twitter users associated, directly or indirectly, with the Islamic State. The data collection included the recovery of social media content generated by ISIS supporters before their accounts were deleted from the Internet and the preservation of online network structures, which I used to estimate ISIS supporters' geographic locations. I used these data to evaluate the extent to which online pro-ISIS radicalization is linked to incidents of anti-Muslim hostility in Europe. I also collected information on the behavior of individuals located in countries targeted by terrorist attacks to examine whether terror acts perpetrated by groups like ISIS systematically provoke anti-Muslims hostility in the West. I find that pro-ISIS extremism is strongly linked to anti-Muslim hostility, and that acts of terrorism systematically increase anti-Muslim sentiment in the West.

To examine what might be done to mitigate these cycles of extremism, I collected data on community engagement events performed in the United States by the Obama Administration, which aimed to increase trust and relationships between the Muslim population and the American government. Using information on the online rhetoric of ISIS supporters in the United States, I show that community engagement events systematically reduced online support for ISIS in areas where they took place. Taken together, these findings suggest that anti-Muslim hostility and pro-ISIS radicalization are strongly linked, and that efforts to stop violent extremism should focus on inclusion and tolerance in the broader population.

1.1 Relationship to existing literature

Understanding why people support violence has long been a focus of conflict research. One branch of this body of work has focused on individual-level drivers of recruitment and support for violence. For example, some have argued that economic or political grievances and strong network ties are important motivators of individual participation in conflict (Gurr, 1970; Horowitz, 1985; Humphreys and Weinstein, 2008; Scacco, 2008; Petersen, 2001). A second, substantial strand of scholarship has theorized on how armed groups mobilize supporters by strategically perpetrating violence against state governments. For example, groups might use terrorism to provoke targeted governments to retaliate indiscriminately in a manner that facilitates support and recruitment (Kydd and Walter, 2006; Crenshaw, 1981; Bueno de Mesquita and Dickson, 2007; Lake, 2002). My research combines these two points of view, which are usually studied separately in the comparative politics and international relations literature, by examining how anti-Muslim hostility might drive support for the Islamic State, on one hand, and how groups like ISIS strategically perpetrate terrorism to manipulate levels of hostility against Muslims, on the other.

My research also relates to literature on immigration and attitudes towards immigrant populations in the West. This body of research examines the economic and cultural factors that facilitate or inhibit immigrant integration in Western countries, emphasizing the powerful influence of natives' attitudes on immigrant integration outcomes (Dancygier and Laitin, 2014; Hainmueller and Hopkins, 2014). Part of this literature has recently studied the integration of Muslim immigrants, empirically documenting the central role of anti-Muslim discrimination in inhibiting successful integration (Adida, Laitin and Valfort, 2014, 2016). Surprisingly, however, this body of work has yet to systematically examine how native attitudes might increase the likelihood of radicalization and violent extremism. My research argues that anti-Muslim sentiment by native populations is an important driver of pro-ISIS radicalization. I examine this proposition in various parts of the dissertation.

Finally, my research contributes to a broader literature on the legacies of violence and terrorism. Research on the psychological effects of terrorism has shown that exposure to terrorist violence increases fear and a sense of threat in targeted population (Huddy et al., 2005; Becker, Rubinstein et al., 2004). Fear and threat, in turn, can strongly intensify animosity between social groups (Quillian, 1995; Stephan and Stephan, 1996; Sullivan, Piereson and Marcus, 1993). Studies examining the effect of 9/11 have shown that the attacks led to an increase in negative attitudes towards Muslim populations, as well as to greater levels of anti-Muslim hate crimes in various Western countries (Coryn, Beale and Myers, 2004; Das et al., 2009; Fekete, 2004; Gould and Klor, 2014). Terrorist attacks have also been shown to increase electoral support for right-wing political parties, who frequently promote of anti-immigration and anti-Muslim policies (Berrebi and Klor, 2008; Kibris, 2011; Getmansky and Zeitzoff, 2014; Fetzer and Soper, 2005; Savage, 2004). My dissertation contributes to this line of work by studying a large number of terrorist attacks in the West, examining how they shaped attitudes towards Muslims and support for far-right parties in Europe and the United States.

1.2 Research approach

To examine the vicious cycle of anti-Muslim hostility and pro-ISIS extremism, I use observational data from multiple sources and several Western countries. As it is difficult to measure radicalization first-hand, I utilized rich observational data on radicalization and extremism located on the social media network known as Twitter. A large majority of the Islamic State's operations is carried out online, so focusing this work on the behavior of Islamic State supporters on Twitter is likely to yield particularly interesting insights. In fact, among over a hundred of individuals charged in the United States with providing material support for ISIS or plotting a violent attack on the organization's behalf, about 62% used social media when they were radicalizing, and among those, 86% expressed their support for ISIS in publicly viewable posts.[1] Twitter allows for the observation of high frequency data that can be analyzed in multiple ways, and I used this data source extensively in the dissertation's chapters.

It is not trivial to capture radical content – for example, this sort of content is often removed by Twitter in response to requests from law enforcement agencies. Thus, I created a data collection algorithm that tracks and documents, in real time, the behavior of over a million accounts of ISIS activists and followers. I exploit new methods in computer science to geo-locate Twitter accounts across the world and match these users to specific locations. This allows me to examine the link between"online" content and the "offline" local context in which these accounts operated.

In the first study, I examine how accounts linked to ISIS were related to areas where anti-Muslim hostility was high. In the second project, I evaluate how terrorist attacks carried out by radical Islamist groups in Europe and the United States spiked anti-Muslim hostility in these countries. In the third paper, I examine how ISIS supporters in the United States reacted to community engagement events aimed to build trust between governments Muslim communities. The combination of multiple sources of observational data across a large number of localities in various countries across the world allowed me to paint a rich picture of the patterns of radicalization and extremism in the West.

[1] See full details in Chapter 2 of the dissertation.

1.3 Summary of the chapters to follow

Chapter 2: From Isolation to Radicalization: Anti-Muslim hostility and Support for ISIS in Europe

In Chapter 2, I focus on pro-ISIS radicalization in Europe. I argue that support for Islamic State in European countries is, at least partly, driven by anti-Muslim hostility. I test this argument using information on the activity of thousands of ISIS activists and the full social network of their followers in France, Germany, Belgium, and the United Kingdom. Matching online radicalization indicators with offline data on vote share for far-right, anti-Muslim political parties, I show that the intensity of anti-Muslim hostility at the local (neighborhood/municipality) level strongly correlates with support for ISIS on Twitter. In substantive terms, I find that an increase of one percentage point in the local-level vote share for far-right parties is associated with a 6% and 3% increase, respectively, in the probability of a user being flagged as ISIS-affiliated and being among the top 1% posters of radical content. In addition, a one percentage-point increase in the right-wing vote share is associated with an average increase of 4,000-9,000 pro-ISIS tweets across the entire sample, including tweets sympathizing with ISIS, expressing anti-West sentiment, and/or interest in foreign fighters and travel to Syria.

As the relationship between pro-ISIS radicalization and support for far-right parties is complex, and may also run in the other direction or be driven by omitted variables, I run several additional tests. First, I take advantage of the high-frequency nature of Twitter data and examine whether events that spur anti-Muslim sentiment, such as terrorist attacks and anti-Muslim protests, are immediately followed by increased posting of pro-ISIS content, especially in areas with high far-right support. Second, I examine whether the results might be driven by the local presence of minority populations. In analyses with data available only in the U.K., I include covariates for the proportion of Muslims, Arabs, Pakistanis, Bangladeshis, and foreign-born in each local area. After controlling for these variables — many of which are negatively or not correlated with radicalization measures — I find that vote share for the far-right remains strongly positively correlated with posting pro-ISIS content on Twitter. Taken together, the findings are consistent with the theory that anti-Muslim animosity increases online radicalization and interest in the Islamic State in these European countries.

Chapter 3: Terrorism as a Provocation Strategy: Transnational Terrorist Attacks and Anti-Muslim Hostility in the West

In Chapter 3, I argue that armed groups strategically use terrorism to manipulate levels of anti-Muslim hostility in the targeted population. I test whether terrorist attacks systematically increase anti-Muslim hostility in the West using data on thirty-six terrorist attacks perpetrated by radical jihadists from 2010 to 2016, and examining how they shaped anti-Muslim attitudes among individuals in targeted countries. This study diverges from past work on the legacies of terrorism, which has focused on aggregate national or sub-national patterns (Gould and Klor, 2014; Hanes and Machin, 2014; Kaushal, Kaestner and Reimers, 2007), by examining the microfoundations of anti-Muslim hostility in the West in studying how terrorist attacks shape the behavior of individual citizens.

For each attack, I collected data on random samples of thousands of individuals in targeted countries, obtaining information on what they posted on Twitter in the two months surrounding the attack. Using text-as-data tools, I created individual-level measures of anti-Muslim sentiment, which were based on the similarity of Twitter posts to an anti-Muslim vocabulary generated from all the tweets that include anti-Muslim hashtags, as well as all tweets generated by politicians from far-right parties who oppose Muslims in Europe.

I find that individuals systematically and significantly increase posting of anti-Muslim content on social media after exposure to terrorism. The effect spikes immediately after attacks, decays over time, but remains significantly higher than pre-attack levels up to a month after the events. The results also reveal that the impact of terrorist attacks on anti-Muslim rhetoric is similar for individuals who already expressed hostility to Muslims before the attacks and those who did not. In other words, terrorism seems to lead individuals in targeted countries to express greater anti-Muslim hostility, regardless of whether they were already hostile before the attacks. Finally, I find that the impact of terrorist attacks on anti-Muslim hostility increases with the lethality of the attack: acts of terror generating more casualties have a stronger effect on anti-Muslim sentiment than attacks generating low number of victims.

Chapter 4: Do Community Engagement Efforts Reduce Extremist Rhetoric on Social Media?

In Chapter 4, I examine what might be done to stop online radicalization and support for ISIS in the West. I focus on community engagement activities in the United States, and examine whether they might have led ISIS sympathizers to changes their pro-ISIS rhetoric on Twitter. Engaging communities in countering radicalization is based on the idea that support for violent extremism can be prevented by involving local communities in efforts to detect early signs of radicalization, and by strengthening trust and relationships between government officials and Muslim communities. While this effort was strongly pursued by the Obama Administration from 2011 to 2016, community engagement initiatives were highly criticized by civil rights and Muslim advocacy groups, who objected the focus on Muslims, arguing that these initiatives suppress the free expression of political opinions in these communities.

In this study, I take a first step at evaluating the possible impact of community engagement events, which are largely understudied in the literature on conflict and violent extremism. I collected data on publicly reported community engagement events held by the Department of Homeland Security's Office of Civil Rights and Civil Liberties (CRCL) from 2014 to 2016 and combined them with high-frequency, geo-located panel data on tens of thousands of individuals in America who follow Islamic State accounts on Twitter. By analyzing over 100 community engagement events in a Difference-in-Differences design, I find that community engagement activities are systematically and significantly associated with a reduction in pro-ISIS rhetoric on Twitter among individuals located in event areas. Specifically, the data show that community engagement events were followed by a decrease in discourse on foreign fighters or travel to Syria, reduction in tweets expressing sympathy with ISIS, and a decrease in the number of tweets discussing the Syrian civil war and life in ISIS-controlled territories. In addition, the observed negative relationship between community engagement activities and pro-ISIS rhetoric was stronger in areas that held a large number of these events.

Chapter 2

From Isolation to Radicalization: Anti-Muslim hostility and Support for ISIS in Europe

Tamar Mitts[1]

[1]I would like to thank Susanne Baltes, Eli Berman, Abhit Bhandari, Christopher Blattman, Jasper Cooper, Lindsay Dolan, Page Fortna, Grant Gordon, Guy Grossman, Kolby Hanson, Macartan Humphreys, Sarah Khan, Sung Eun Kim, Egor Lazarev, Summer Lindsey, Joshua Mitts, Suresh Naidu, Jacob Shapiro, Kunaal Sharma, Jack Snyder, Camille Strauss-Kahn, and Lauren Young for their advice and feedback on various stages of this project. This research was approved by the Columbia University Institutional Review Board under protocol IRB-AAAQ7657.

Abstract

What explains online radicalization and support for ISIS in the West? Over the past few years, thousands of individuals have radicalized by consuming extremist content online, many of whom eventually traveled overseas to join the Islamic State. This study examines whether anti-Muslim hostility might drive pro-ISIS radicalization in Europe. Using new geo-referenced data on the online behavior of thousands of Islamic State supporters in France, the United Kingdom, Germany, and Belgium, I study whether the intensity of anti-Muslim hostility at the local (neighborhood/municipality) level is linked to pro-ISIS radicalization on Twitter. Results show that local-level measures of anti-Muslim animosity correlate significantly and substantively with indicators of online radicalization, including posting tweets sympathizing with ISIS, describing life in ISIS-controlled territories, discussing foreign fighters, and expressing anti-West sentiment. High-frequency data surrounding events that stir anti-Muslim hostility – terrorist attacks and anti-Muslim protests in Europe – show the same pattern.

2.1 Introduction

Since 2011, about 30,000 foreign fighters traveled to Syria and Iraq to join the efforts of the Islamic State (ISIS), according to a recent assessment by US intelligence analysts (Schmitt and Sengupta, 2015). Fighters travel to ISIS territories from all over the world, and high numbers come from Western countries, such as France, Britain, Belgium, Germany and the United States. A very large proportion of Western recruits are being radicalized online by consuming extremist content on the Internet and social media (Carter, Maher and Neumann, 2014; Vidino and Hughes, 2015*b*). Radicalization and support for ISIS are not limited to certain social groups or specific national grievances; instead, radicalized individuals come from many different backgrounds, age groups, education, and income levels (Greenberg, 2016). What motivates Westerners to radicalize and support groups like the Islamic State? How can an organization attract so many individuals to a conflict not their own?

This study brings together research on violent extremism and radicalization, along with literature on immigration in the West, to examine how anti-Muslim sentiment is linked to radicalization and support for the Islamic State in European countries. I argue that hostility towards Muslims in the West can lead individuals to seek comfort and acceptance elsewhere, making radical messages promulgated by foreign rebels seem attractive. A large body of research on immigration to the West studies factors that facilitate or inhibit immigrant integration, with a particular focus on economic outcomes (Dancygier and Laitin, 2014). This literature emphasizes the powerful role that natives' attitudes play in this context, and points to cultural, economic, and psychological factors that determine natives' acceptance, or lack of acceptance, of immigrants in social and economic settings (Hainmueller and Hopkins, 2014).

A recent strand of this important body of work has focused on discrimination against

Muslim immigrants in particular, empirically documenting the central role of anti-Muslim discrimination in facilitating Muslims' lack of integration. In France, for example, Adida, Laitin and Valfort (2014, 2016) found that Muslims and non-Muslims are stuck in a vicious cycle in which the latter discriminate against the former, falsely equating "Muslim" and "Jihadist," and Muslims, in turn, tend to distrust non-Muslims and withdraw from French society, thus perpetuating their non-assimilation. But this body of research has yet to examine other outcomes of discrimination. Focusing primarily on social and economic integration, it has not systematically considered how native attitudes towards immigrants might increase the likelihood of foreign fighter radicalization.

One of the most distinctive aspects of the Islamic State's recruitment strategies is its extensive use of social media. The organization not only distributes provocative content to general audiences on the Internet, it uses social networks on *Twitter*, *Facebook*, and related platforms to attract new members from all over the world. Twitter has been widely used by the organization, because it provides technical advantages such as large-scale public dissemination of content (Klausen, 2015). Studies documenting the usage of Twitter by Western foreign fighters have noted that it played a central role in their radicalization process by intensifying their mental and emotional connection to the war events on the ground (Carter, Maher and Neumann, 2014). Potential recruits find it appealing to connect to the organization through Twitter, as the platform enables the anonymous consumption of radical and extremist ideas, without being exposed to the risk of physically interacting with a recruiter (Berger, 2015). In fact, the organization's online radicalization operation is so vast and extensive that many security agencies find it challenging to keep track of every aspect of these activities (Homeland Security Committee, 2015).

In this study, I take advantage of the presence of this widespread online radicalization, and the availability of large amounts of public Twitter data, to examine whether anti-Muslim hostility is linked to support for ISIS in Europe. Using an original method, I collected granular data on the social media activity of about 15,000 accounts of ISIS activists, as well as the full social network of their followers across the world ($N \approx 1.6$ million). I monitored the online behavior of ISIS activists and their followers in real-time, capturing their activity prior to account suspension, and recorded textual and image content, which I use for analysis.

Using computer science methods to predict the physical geographic location of Twitter users, I matched user content to local-level administrative data from the four European countries with the highest share of Western foreign fighters: France, the United Kingdom, Germany, and Belgium (Barrett et al., 2015). I collected data on levels of unemployment, the share of immigrants and asylum seekers in each locality, and local-level vote share for far-right, anti-Muslim parties in recent elections across Europe. As voting for far-right parties strongly correlates with anti-Muslim sentiment,[2] I use vote share for these parties as a local-level measure of anti-Muslim hostility, examining whether it predicts support for ISIS on social media.

I developed several measures of online radicalization and support for ISIS on Twitter. Using supervised machine learning, I classified millions of tweets in English, Arabic, French, and German along various dimensions of ISIS support. These include expressing sympathy with ISIS, tweeting about the life of fighters in ISIS-controlled territories, expressing an interest in traveling to Syria or becoming foreign fighters, and generating anti-West content. In addition, I kept track of which users were flagged as ISIS activists by several hacktivist groups, and also noted when they were suspended from Twitter. I use each of these to measure pro-ISIS radicalization.

Results show that local-level vote share for far-right, anti-Muslim parties in France, the United Kingdom, Germany, and Belgium is a significant predictor of online radicalization. In substantive terms, an increase of one percentage point in the local-level vote share for

[2]See more information in section 2.2.

far-right parties is associated with a 6% and 3% increase, respectively, in the probability of a user being flagged as ISIS-affiliated and being among the top 1% posters of radical content. A one percentage-point increase in the right-wing vote share is associated with an average increase of 4,000-9,000 pro-ISIS tweets across the entire sample, including tweets sympathizing with ISIS, expressing anti-West sentiment, and/or interest in foreign fighters and travel to Syria.

As the relationship between online radicalization and support for far-right parties is complex, and may also run in the other direction or be driven by omitted variables, I run several additional tests. First, I take advantage of the high-frequency nature of Twitter data and examine whether events that spur anti-Muslim sentiment, such as terrorist attacks and anti-Muslim protests, are immediately followed by increased posting of pro-ISIS content, especially in areas with high far-right support. Second, I examine whether the results might be driven by the local presence of minority populations. In analyses with data available only in the UK, I include covariates for the proportion of Muslims, Arabs, Pakistanis, Bangladeshis, and foreign-born in each local area. After controlling for these covariates—many of which are negatively or not correlated with radicalization measures—I find that vote share for the far-right remains strongly positively correlated with posting pro-ISIS content on Twitter. Taken together, the findings are consistent with the theory that anti-Muslim animosity increases online radicalization and interest in the Islamic State in the West.

2.2 Anti-Muslim hostility and radicalization in the West

In this Section, I offer a theoretical framework whereby radicalization emerges in a vicious cycle of reacting to, and feeding, anti-Muslim hostility in the West. In the subsections that follow, I situate the argument of this paper within this vicious cycle. In Subsection 2.2.2, I suggest that anti-Muslim hostility is reflected in support for far-right parties in Europe.

In Subsection 2.2.3, I show that this hostility leads to two forms of Islamist[3] radicalization: socially visible acts of extremism (e.g., terrorism) and non-socially visible actions (e.g., consuming extremist content on the Internet). I argue that the less-noticeable nature of the latter makes it possible to evaluate whether the empirical evidence supports the view that animosity against Muslims leads individuals to express support for ISIS.

2.2.1 Radicalization

> *"You know what I hate about the West? ... They call everyone who made hijrah to Shaam [Syria] traitors to their country. How can you be a traitor to your country when it never fully accepted you or your beliefs? Today, my first day of [university] I got called a "suicide bomber" and was the butt of several ISIS jokes. All because I chose to dress modestly. May Allah bless you and everyone who went there and all those who are trying hard to make his word one. It really sucks being stuck here."*
>
> - Muslim woman living in a Western country[4]

> *"When I am bored I make a map of what would be a united Muslim caliphate, and when it comes to [where I live]"*
>
> - French ISIS supporter, June 2015.[5]

Why do individuals living in Western countries begin to support the Islamic State? What attracts people to ISIS' extremist ideology? Hoda Muthana, an American student from Alabama, was radicalized on social media after opening a secret Twitter account without her parents' knowledge. After interacting with ISIS supporters on Twitter, she adopted radical interpretations of Islam and eventually traveled to Syria to join the organization (Hall, 2015). Ali Shukri Amin, an American teenager from Virginia, found solace from his troubled life in the virtual communities of ISIS activists on Twitter. In the end, Amin

[3] In this paper, I use the term 'Islamist' when an referring to ideology that advocates or supports of Islamic militancy or fundamentalism.

[4] Source: Pooley (2015) Her note was posted on the *Ask.fm* account of an ISIS foreign fighter.

[5] Original post written in French. Source: the author's database on content produced by ISIS supporters in the West.

disconnected from his family and friends, spread ISIS propaganda to thousands of followers online, and recruited one of his friends to travel to Syria to become a foreign fighter (Shane, Apuzzo and Schmitt, 2015; Robinson, 2015). Among over a hundred individuals charged in the United States with providing material support for ISIS or plotting a violent attack on the organization's behalf, about 62% used social media when they were radicalizing, and among those, 86% expressed their support for ISIS in publicly viewable posts.[6] This pattern poses a puzzle: how was ISIS able to inspire the radicalization of so many people in the West?

A large literature has sought to explain the causes of radicalization and violent extremism, especially in the context of militant Jihad. Most agree that radicalization involves a change in ideology or beliefs that support indiscriminate violence against civilians for political reasons, or a group that represents this ideology and actions (Borum, 2011; Crossett and Spitaletta, 2010; McCauley and Moskalenko, 2008; Wilner and Dubouloz, 2010; Sedgwick, 2010).[7] Earlier work sought to identify a 'terrorist personality type,' (Horgan, 2008) but subsequent research found little evidence supporting this idea, and argued that radicalization is a process where contextual factors, such as social discrimination, political repression, or economic crises, prompt individuals to open up to radical worldviews (Borum, 2011). In this paper, I argue that a hostile, anti-Muslim environment can trigger this process of pro-ISIS radicalization.

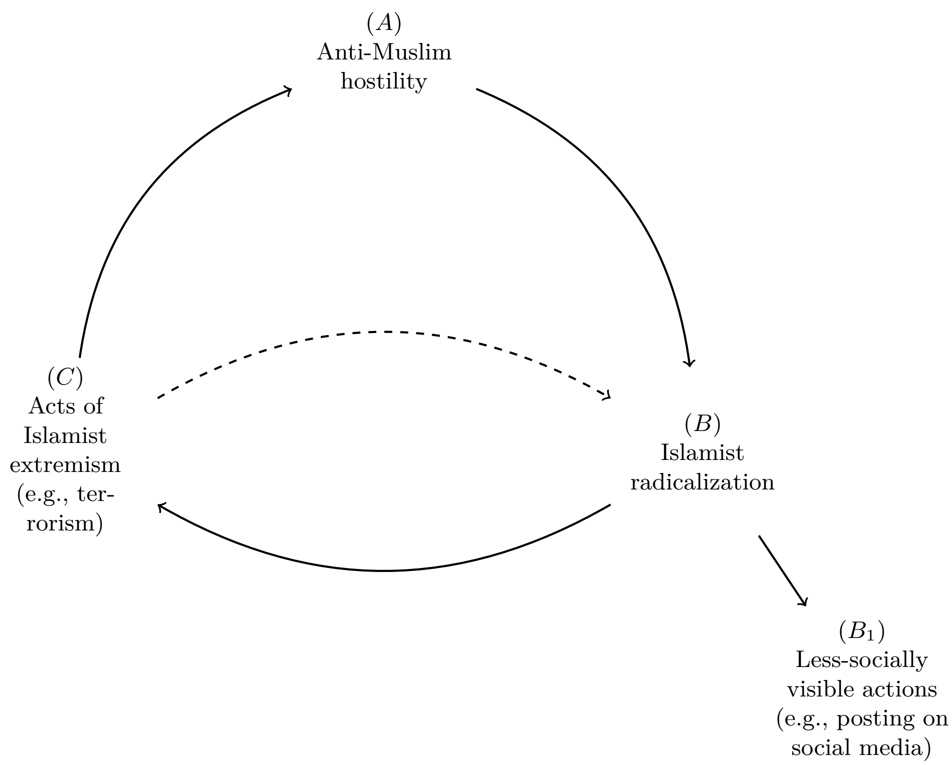
Of course, Islamist radicalization does not occur in isolation. Radicalization is part of a vicious cycle which reacts to, and feeds, anti-Muslim hostility in the West. Figure 2.1 illustrates this endogenous system, in which anti-Muslim hostility (node A) leads to Islamist radicalization (node B). In return, highly visible actions by Islamist extremists (node C)

[6] These data were collected by the author, and come from FBI investigations of individuals in the United States. See full details in the Appendix.

[7] In this study, I use the term 'radicalization' more narrowly, i.e., to refer simply to expressions of support for and sympathy with the Islamic State—a group that represents an extreme ideology and carries out indiscriminate violence against civilians for political reasons.

drive further anti-Muslim hostility (node A). As I explain further below, radicalization can also take the form of actions that are not always socially noticeable (B_1), which do not easily feed back into the vicious cycle.

Figure 2.1: A vicious cycle of radicalization and anti-Muslim hostility



2.2.2 Anti-Muslim hostility and support for far-right parties

Animosity against Muslims in the West has been rising in recent years, especially after 9/11 (Stack, 2015; Burrows, 2016). Examples include setting fire to mosques, spreading anti-Muslim graffiti, and physically attacking individuals who practice Islam. Take the case of Ms. Khola Hasan, an Islamic scholar from the U.K.'s Epping Forest region, who has been targeted by anti-Muslim violence multiple times in recent years. In an interview with *The Guardian*, she said, "I was walking down Epping High Street and a man shouted at me 'You bloody ISIS supporter.' Another time ... someone stopped their car and threw an empty glass bottle at me. I was absolutely terrified." (Flaig, 2016)

Epping Forest is among the constituencies with the highest vote share for far-right parties in the United Kingdom. In the 2015 general elections, over 18% of its voters voted for far-right parties, putting the locality at the top 10% of far-right vote share in the country. A similar pattern is observed in other European localities with high far-right support. In Dartford, U.K., right-wing activists recently launched an "anti-halal operation," targeting Muslim restaurants selling halal food with the claim that they support terrorism by paying a zakat religious tax (Kent Online, 2015). In Provins, France, where vote share for the Front National party in the 2015 Departmental Elections was above 37%, a local mosque was recently desecrated with anti-Muslim graffiti (Inge, 2013).

Indeed, far-right parties are one of the most prominent mobilizers of anti-Muslim sentiment in contemporary Europe. A common theme in the platforms of these parties is support for exclusionary, "nativist" populism that combines nationalism and xenophobia, seeking to ostracize groups with certain cultural, religious, or ethnic characteristics (Golder, 2016). For example, France's Front National party has long blamed Muslim immigrants for many of the country's social problems, ranging from unemployment to security and national unity (Adida, Laitin and Valfort, 2016; Front National, 2016).[8]

Several scholars have suggested that far-right voting is strongly linked to anti-Muslim sentiment (Lubbers and Scheepers, 2002; Norris, 2005; Rydgren, 2008). Using data from the European Social Survey Round 7 Data (2014) (ESS Round 7, 2014), I tested the relationship

[8] While the overall popularity of far-right, anti-Muslim parties in Europe has increased nationally, support for these parties still varies significantly at the local level. Figure 2.24 in the Supplementary Information materials shows the vote share for far-right parties in France, Germany, and the U.K. at the electoral constituency level in national elections taking place in recent years. Research on support for far-right parties in has shown that it tends to be stronger in areas where minority communities form a smaller share of the population (Biggs and Knauss, 2012). In this paper, I use local-level data on vote share for far-right parties as a proxy for local anti-Muslim hostility, and examine whether it is linked to online radicalization and support for ISIS among potential recruits.

between far-right voting and anti-Muslim attitudes in Europe. Table 2.1 shows that there is a strong correlation between holding anti-Muslim and anti-immigrant attitudes and self-identifying as a far-right supporter (Panel A) or voting for far-right parties (Panel B).[9] The regressions control for a large number of demographic variables that might also explain anti-Muslim attitudes, such as being native-born, education, income, gender, age, and religious beliefs.

These patterns suggest that local-level support for far-right parties can reflect an atmosphere of anti-Muslim hostility. I use sub-national data on vote share for far-right parties to proxy for local anti-Muslim animosity, and examine whether it is linked to online radicalization and support for ISIS among potential recruits located in these areas. In the following section, I explain how I created measures for online support for ISIS by identifying and observing in real-time the content and social media activity of individuals at risk of radicalization.

2.2.3 The social visibility of Islamist radicalization

In this Subsection, I draw a distinction between two ways that radicalization can be expressed.[10] One form of radicalization consists of what I refer to as "socially visible" actions, which are those that have a strong, prominent effect on an individual's environment and surroundings. An obvious example is the perpetration of terrorist attacks. Much of the literature on radicalization has focused on this publicly salient form (Wilner and Dubouloz, 2010).

However, radicalization reflects an internal psychological process that does not always

[9]The parties used to create the voting measure for the far-right are Front National in France, United Kingdom Independence Party (UKIP) in the United Kingdom, National Democratic Party of Germany (NPD) and Alternative for Germany (AfD) in Germany, and Vlaams Belang in Belgium.

[10]Of course, this is not a true dichotomy, and many forms of radicalization will lie in between these extremes. But this is a useful way to identify types of behavior that are more or less likely to feed back into the vicious cycle.

Table 2.1: Far-right support and anti-Muslim attitudes in Europe

	(1) Do not allow Muslims in country	(2) Disapprove immigration of different race/ethnic groups	(3) Disapprove relative marrying someone from a minority race/ethnic group	(4) Do not want a boss from a minority race/ethnic group	(5) Immigrants make crime worse
	A. Far-right self placement				
Far-right self placement	0.12***	0.37***	0.99***	0.39	0.41**
	(0.03)	(0.07)	(0.27)	(0.24)	(0.18)
Constant	0.05	2.14***	2.09***	1.58***	6.78***
	(0.04)	(0.10)	(0.37)	(0.32)	(0.28)
Demographic controls	✓	✓	✓	✓	✓
R^2	0.054	0.075	0.068	0.075	0.023
Observations	3,850	3,874	3,894	3,867	3,837
	B. Far-right voting				
Voted for far-right party	0.26***	0.65***	1.91***	1.49***	1.23***
	(0.05)	(0.09)	(0.34)	(0.35)	(0.24)
Constant	0.06	2.15***	2.12***	1.56***	6.77***
	(0.04)	(0.09)	(0.37)	(0.32)	(0.28)
Demographic controls	✓	✓	✓	✓	✓
R^2	0.070	0.085	0.076	0.084	0.033
Observations	3,850	3,874	3,894	3,867	3,837

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: The table reports the correlations between voting for far-right parties in France, Belgium, Germany, and the United Kingdom and holding anti-Muslim and anti-immigrant attitudes. The *Far-right self placement* variable is an indicator coded 1 for individuals who identify as '10' (farthest on the right) on a 1-10 scale of left-right placement. *Voted for far-right party* is an indicator variable coded 1 for individuals who voted for Front National (FN) in France, United Kingdom Independence Party (UKIP) in the United Kingdom, National Democratic Party of Germany (NPD) and Alternative for Germany (AfD) in Germany, and Vlaams Belang (VB) in Belgium. The table presents estimates from ordinary least squares regressions of the outcome variables reported in columns (1) through (5) on indicators of support for far-right parties, controlling for being native-born, education, income, gender, age, and religion. Data source: European Social Survey Round 7 Data (2014).

culminate in visible provocative actions. A study of the radicalization process of Islamist activists in the United Kingdom found that it began with a "cognitive opening" to extremist Islamic ideas, followed by immersion in radical ideology (Wiktorowicz, 2005). Crucially, while this process of personal radicalization can sometimes lead individuals to carry out actions that have a high degree of social visibility, many times they do not.[11]

In other words, it is possible to identify behaviors that reflect a process of radicalization but are largely hidden from an individual's social environment. For example, ISIS supporters may post expressions of sympathy with the organization on social media, but these posts are not necessarily visible to people in their immediate surroundings or geographic neighborhood. In several accounts of individuals who radicalized in the United States, support for extremist ideology was only expressed in private messages sent between these individuals and undercover FBI agents (Jacobs, 2014; King, 2016). That's not to say that these sorts of less socially-visible actions will never be noticeable by one's society, but their infrequent, highly personal nature makes it unlikely that they will intensely drive the vicious cycle, especially when measured at a disaggregated level in time and space.

For these reasons, this sort of less noticeable, personal radicalization is likely to be more amenable to empirical analysis than highly visible forms of extremism (e.g., terrorist attacks) that raise more fundamental endogeneity concerns. In this paper, I consider whether the empirical evidence is consistent with the theory that anti-Muslim hostility leads individuals to express support for ISIS. In the following Section, I explain how I measured these sorts of less-socially visible outcomes by identifying and observing in real-time the content and social media activity of individuals at risk of radicalization. I also discuss aspects of my empirical research that are designed to isolate the link between anti-Muslim hostility and online radicalization.

[11] Empirically, I find that there is little geographic link between terrorist attacks and empirical measures of radicalization; in fact, the two tend to be weakly negatively correlated. See Appendix for more information.

2.3 Data

In this Section, I describe the original data used in this study. In Subsection 2.3.1, I explain how I identify ISIS activist and follower accounts on Twitter—a nontrivial challenge akin to finding a needle in a haystack. In Subsection 2.3.2, I explain the algorithm that I use to predict the physical geographic locations of Twitter users in my dataset. In Subsection 2.3.3, I define the measures of online radicalization, which serve as the outcome variables in this study, and provide summary statistics on these variables. Finally, in Subsection 2.3.4, I describe the local-level administrative data which are matched to each user on a geographic basis, and thus serve as the independent variables in this study.

2.3.1 Identifying ISIS activist and follower accounts on Twitter

Finding people on Twitter who are sympathetic to the Islamic State and/or possibly radicalizing is a challenging task. While there are many ISIS accounts active on Twitter—some have estimated as many as 40,000–125,000 (Berger and Morgan, 2015; Isaac, 2016)—this is still a small figure compared to the overall number of Twitter users, which amounts to about 319 million.[12] In a way, it is like finding a needle in a haystack. Previous methods for identifying ISIS supporters on Twitter include measuring the networks of selected seed accounts (Berger and Morgan, 2015; Chatfield, Reddick and Brajawidagda, 2015) and qualitatively analyzing a handful of accounts of recruits known to have migrated to Syria and Iraq (Carter, Maher and Neumann, 2014; Pooley, 2015).

In this project, I identify ISIS accounts by tracking in real-time lists published by several anti-ISIS hacking groups. As the organization's activity on social media intensified in the past several years, groups such as Anonymous and Controlling Section (@CtrlSec) began monitoring social media accounts identified with the organization and publicly flagging

[12] As of August 2016. `http://www.statista.com/statistics/282087/number-of-monthly-active-twitter-users/`

them for suspension. At the beginning of 2015, the group @CtrlSec asked social media users to help find ISIS accounts on Twitter (see Figure 2.19 in the Appendix), an effort that led to the suspension of thousands of accounts in a matter of days. Since then, the monitoring, flagging, and suspension of ISIS accounts continues,[13] and this dissertation project leverages this information to identify ISIS activists' accounts.

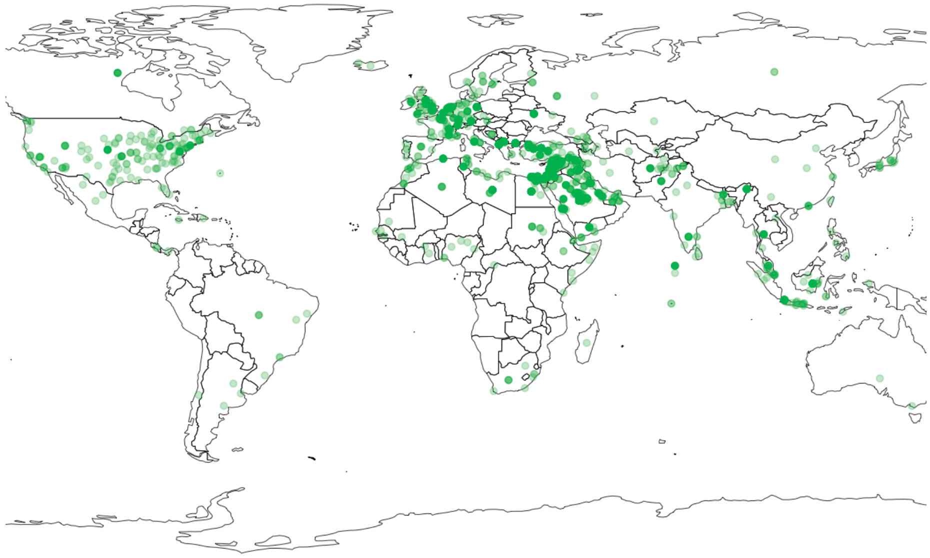
I designed an algorithm that since December 2015 has been continually monitoring and recording Islamic State accounts identified by @CtrlSec (see Figures 2.21 and 2.22 in the Appendix for examples of such accounts). Immediately upon observing a new account in the @CtrlSec list, I downloaded the complete historical "timeline" of tweets for the account, as well as its user profile, which includes self-described location, time zone, and list of the account's friends and followers.[14] This real-time data collection enabled me to capture information on accounts of about 15,000 ISIS activists before they were deleted from the Internet. In addition, I collected data on the online activity of about 1.6 million followers of these @CtrlSec-identified ISIS accounts, to identify a set of people "at risk" of becoming ISIS activists, as well as about 450,000 friends of these followers. The database, which is described in detail in the Appendix, contains user-level information, taken as "snapshots" of each user's profile at various points in time, as well as tweet-level information on over 61 million tweets, collected from all ISIS activist accounts, as well as a subset of the followers. Figure 2.2 shows the geographic distribution of the accounts of ISIS activists across the world.

---

[13] Earlier in 2015, Twitter announced that it has suspended about 125,000 ISIS accounts, many of which are believed to be flagged by @CtrlSec. See: `http://www.nytimes.com/2016/02/06/technology/twitter-account-suspensions-terrorism.html?_r=0`; as well as: `http://www.theatlantic.com/international/archive/2015/10/anonymous-activists-isis-twitter/409312/`

[14] On Twitter, a "friend" of account i is an account that i follows.

Figure 2.2: Predicted locations of ISIS activists on Twitter across the world



Note: The figure plots the predicted locations of accounts flagged as ISIS activists by @CtrlSec. Locations are predicted using a recent method in the computer science literature—spatial label propagation—to predict the geocoordinates of social media users (Jurgens, 2013; Jurgens et al., 2015).

2.3.2 Predicting geographic locations of ISIS activists and followers

A central aspect of this project involves predicting the geographic location of Islamic State activists and followers on Twitter, in order to match them to geographic data on socio-economic variables that might predict online radicalization. Since a very small share of Twitter users enable geo-tagging of their tweets or provide location information in their accounts,[15] social network and computer science researchers have developed methods in recent years to triangulate a user's location based on locations provided by their networks of friends and followers (Backstrom, Sun and Marlow, 2010; McGee, Caverlee and Cheng, 2013; Jurgens et al., 2015). I employ a spatial label propagation algorithm developed by Jurgens (2013) to predict Twitter users' locations, which performs three rounds of prediction to maximize predictive accuracy.

[15]For example, in my database, only 26% of the users enable geo-tagging of their tweets, and 37% provide self-described location information.

Spatial label propagation algorithms rely on the finding in social network research that location information in a user's online network is a powerful predictor of a user's offline geographic location (Goldenberg and Levy, 2009; Takhteyev, Gruzd and Wellman, 2012; McGee, Caverlee and Cheng, 2011). While social media platforms allow people to connect with others across the globe, recent studies have found that physical relationships in the offline world still strongly influence online social relationships. When people live their lives offline, they form relationships that subsequently transfer to the online world—e.g., co-workers or classmates who meet offline and then connect on social media platforms. As a result, a large share of individuals' online social network usually includes geographically close friends. Figure 2.25 in the Appendix, which is taken from Jurgens (2013), shows that across various social networks on different platforms, the majority of individuals in the network had at least one friend that was located within 4 kilometers. The Appendix provides more information on the details of this location prediction process, as well as a discussion of its out-of-sample predictive accuracy.

While this method is imperfect and subject to prediction error, the rich data that it provides allows us to examine the local-level correlates of online support for Islamic State in Europe. As existing quantitative research on ISIS foreign fighter recruitment has so far remained at the country level, this is an important step forward. In addition, while prediction errors make estimations more noisy, there is little reason to think they are plagued by systematic biases.[16] Location predictions are carried out on a very large and relatively deep network of over 2 million Twitter users across the world. Location prediction errors can bias the results only if they affect the *network* structure of individuals showing support for ISIS, e.g., by leading them to *strategically* choose friends so that their locations are systematically predicted (incorrectly) in areas with higher vote share for far-right parties. Strategic choice of friends in this way is difficult to perform systematically.

[16] A test of the correlation between the prediction errors and far-right vote share shows no systematic relationship. The results of this test are reported in the Appendix.

Moreover, location prediction is carried out for *all* users in the database and analysis is carried out across thousands of localities in four countries. For systematic biases to be present, location predictions for ISIS supporters would have to appear systematically across countries in a pattern that correlates with far-right party vote-share locally. To address the concern that Internet usage varies across rural and urban areas, regressions control for local population size.

2.3.3 Measuring online radicalization

I measure online radicalization using various user-level and tweet-level variables from the ISIS activists/followers database. Since my pool of subjects consists of individuals who already follow one or more ISIS accounts, my analysis is limited to people who already show signs of interest in the organization. Nonetheless, since the followers sample consists of a range of accounts—from individuals who are already ISIS activists to accounts that are countering ISIS—being included as a follower in the database does not imply that one is actually radicalizing. To address the concern that I measure anti-ISIS accounts as 'radicalized,' I use textual and social network information to find tweets and users who are more likely to reflect pro-ISIS content.

Online radicalization measures are constructed as follows. First, I employ data from user-level fields to create indicators for whether a given user is flagged as an ISIS activist by @CtrlSec. Second, I use data on account suspension to code whether a user is suspended from Twitter for being associated with ISIS. Third, I use the network information in the database to count the number of ISIS accounts that each user follows. Fourth, I create textual measures for the number of pro-ISIS tweets posted by each user along several dimensions of ISIS support.

To generate the textual outcomes, I use supervised machine learning to classify tweets in English, Arabic, French, and German into one or more of these categories:

1. *Sympathy with ISIS* - expressions of support or sympathy with the Islamic State, its ideology and its activities in territories under its control
2. *Life in ISIS territories* - tweets describing the life of ISIS activists in the territories controlled by the Islamic State
3. *Travel to Syria or foreign fighters* - tweets describing interest or intent to travel to Syria, and/or discussion of foreign fighters
4. *Syrian war* - tweets describing events in the Syrian civil war and/or discussion/analysis of those events
5. *Anti-West* - anti-West rhetoric, criticizing Western countries' foreign policy and military operations in the Middle East
6. *Islam* - expressions of faith in the Islamic religion, Islamic quotes, and prayers and/or requests for prayers

While some of these topics may not signal online radicalization—for example, tweets expressing faith in Islam or tweets criticizing the West are likely to be completely benign in most cases—the combination of these with other topics that more directly reflect support for ISIS, can plausibly capture radical content. Therefore, I measure online radicalization by looking at the distribution of tweets across all of these topics combined, coding Twitter users who are on the higher end of the distribution of posting on these topics as more strongly supporting the organization. Table 2.2 shows examples of English language tweets for each of these topics.

The supervised learning process works as follows. First, human coders from two crowdsourcing platforms, Amazon Mechanical Turk and Crowdflower, labeled a random sample of posts by hand.[17] Then, an algorithm used information on the words in each labeled post to "learn" the categorization rules and classify unlabled posts.[18] I obtained a random sample of tweets posted by ISIS activists in English, Arabic, French, and German to create

[17]See Figure 2.26 in the Appendix for an example of instructions for the classification task in the Crowdflower platform.

[18]See Grimmer and Stewart (2013) for a review and more information on supervised machine learning methods to classify text, and James et al. (2013) for an introduction to machine learning in general. The Appendix provides more details on the supervised learning method used in this paper.

Table 2.2: Examples of tweets in different topics

Sympathy with ISIS
Jihad is the greatest of all deeds #IslamicState
Show everything from the Islamic State and other groups in Syria. It's important to hear all sides of the story.
Assalam o Alaikom to All Islamic State Brothers
In sha Allah we will have honor again #IslamicState
Life in ISIS territories
#Aljazeerah reports from inside the city of #Raqqa and shows how the #IslamicState runs the daily life
English Testimony from a girl in #Yarmouk Camp about the #IslamicState
The glorious and mighty army of the Caliphate: Young kids ready to blow themselves up.
Health services in Islamic state
Wedding of an #ISIS fighter in #Raqqa:
In Ribaat... Nice feeling really; Sit with the bros, drink tea, read Quran, relax & just observe the enemy! #Syria
Travel to Syria or foreign fighters
a lot of foreign fighters still coming in. Seems a lot responding to the call of the scholars of General March, also indicating open way in!
is the door to sham open?
i want your kik please akh, maybe there is sister in my country that have more money and looking for hijra too...
come on join us at syam..
Dutch fighters in ar-Raqqa province #Syria
Syrian war
#IS fighters readying to fight an invasion of Yarmouk Camp by Assad's allies Jaysh Al-Islam and Liwa Sham Al-Rasool
Massive destruction in Douma today after one of Assad's almost daily air strikes on the city. #Syria #Damascus
#Syria - The evil #Assad regime lost Busra al-Harir so they tortured a 6 year old girl out of revenge...
Massive explosion rocked entire of #Ramadi city. No further details yet.. #Iraq #ISIS
Anti-West
America has been at war 222 out of 239 years since 1776. Let that sink for a moment.
If Islamic State terror is evil why would Western State war be good?
US-led wars on terror have killed four million Muslims since 1990
It's sad when I am more afraid of our government then #ISIS ! At least I know #ISIS hates #America #Government =wolves
Why are we shocked at ISIS brutality but not shocked by US British & European brutality?
Islam
Call upon Me; I will respond to you. #Quran 40:60
My identity is in who Allah says I am not in who others say I am. Allah's opinion is the only one that truly matters.
To think that Allah Almighty is present with you at every given moment is the most excellent form of #faith.
The beauty of Sujood is such that you whisper silently in to the ground and it's heard up in the Heavens.
May allah bless you brother.......

a training set for the classification model.[19] Each tweet was labeled by three coders, and and label(s) were retained for a given tweet only if at least two out of the three coders

[19]English, Arabic, French and German are used in 76% of the tweets in the database. As the proportion of tweets in the database varies by language, the size of the training set accordingly varies for different languages: English ($N = 9,926$), Arabic ($N = 10,631$), French ($N = 6,158$), and German ($N = 3,011$).

assigned the same label(s) to the tweet.

Since Twitter textual data are very noisy, and radical pro-ISIS content is rare, many tweets in the database were coded as unrelated to any of the above categories.[20] To facilitate statistical prediction, I follow King and Zeng (2001) and randomly over-sample pro-ISIS tweets and randomly under-sample unrelated tweets to obtain a class proportion of 0.5 for each of the categories, for each topic, for each language. I trained separate logit models using the labeled rebalanced training sets for each category in each language. For all specifications, I used the the elastic-net generalized linear model (Friedman, Hastie and Tibshirani, 2010), selecting the regularization parameter λ by cross-validation to maximize the area under the ROC curve. Using this method, the models were able to predict radicalized content with an in-sample accuracy of above 95%. More metrics on the performance of the models for each topic and language are reported in Section 2.6.3 in the Appendix. The classification models for each topic and language were then employed on the full set of tweets in the database to classify each unlabeled tweet as belonging to one or more of these categories. Appendix Table 2.36 shows the top 50 words for each topic in English language tweets.

To measure users' posting of radicalized pro-ISIS content, I counted the number of tweets classified in these six categories for each user. I also created a combined measure that counted the number of tweets falling into any of these categories. To ensure that I capture users that post highly pro–ISIS content, I created an indicator that is coded 1 for users who are at the top 1% of the distribution of radicalized content posting and 0 otherwise.[21] Panel A in Table 2.3 provides summary statistics for these various measures of online radicalization.

[20]See Appendix for details on the classes for each outcome and language.

[21]I chose this cutoff in order to be conservative and not erroneously classify as radicalized individuals who post less radical content. As reported in Table 2.27 in Section 2.6.7 in the Appendix, results hold in estimations with cutoffs using top 5%, 10% 15%, and 20%.

Table 2.3: Summary statistics

Statistic	N	Mean	St. Dev.	Min	Max
A. Dependent variables					
Tweets with pro-ISIS content (#)	174,479	35.304	90.508	0	1,933
Pro-ISIS content (top 1%)	174,479	0.010	0.099	0	1
Tweets with pro-ISIS content, excluding Islam (#)	174,479	25.712	66.411	0	1,421
Pro-ISIS content, excluding Islam (top 1%)	174,479	0.010	0.099	0	1
Sympathy with ISIS (#)	174,479	4.657	12.297	0	277
Life in ISIS territories (#)	174,479	6.263	16.522	0	308
Travel to Syria or foreign fighters (#)	174,479	6.645	17.195	0	343
Syrian war (#)	174,479	3.707	10.016	0	251
Anti-West (#)	174,479	4.440	11.982	0	287
Islam (#)	174,479	9.592	24.594	0	512
Number of ISIS accounts following	174,474	5.452	23.863	0	3,216
Flagged as an ISIS activist	174,479	0.005	0.070	0	1
Suspended by Twitter	174,474	0.041	0.199	0	1
B. Independent variables					
Far-right vote share (%, local level)	116,493	13.208	9.026	0	53.805
Unemployed (%, local level)	170,654	5.124	2.410	0	41
Immigrants unemployed (%, local level)	90,521	1.889	0.993	0	9
Foreigners/non-citizens (%)	171,077	10.466	7.405	0	89.026
Population	171,548	815,078	1,078,663	3	3,292,365

Note: This table reports summary statistics for the sample of ISIS activist and followers who are predicted to be located in France, Germany, Belgium, and the United Kingdom.

2.3.4 Independent variables

To measure anti-Muslim hostility at the local level, I created variables for the vote share for far-right parties at the electoral constituency level in France, Germany, Belgium, and the United Kingdom. I downloaded data on official election results in these countries, and calculated the percent of votes for parties associated with far-right positions. Table 2.4 shows the elections and parties used to construct this variable. Using Twitter users' predicted geo-location data (see Section 2.3.2 for more details), I matched users in my database to electoral constituencies, thereby assigning users to different areas with varying degrees of far-right support. Panel B in Table 2.3 shows that vote share for these parties varies substantially, where some users are located in areas with zero vote share for far-right parties, and others in areas with more than 50% support for these parties.

Table 2.4: Far-right parties in recent European elections

Country	**Election**	**Far-right parties**
France	2015 departmental elections	Front National (FN)
Germany	2013 Federal elections	National Democratic Party of Germany (NPD); Alternative for Germany (AfD)
United Kingdom	2015 general elections	British Democrats; British National Party; Liberty GB party; National Front party; United Kingdom Independence Party (UKIP)
Belgium	2014 Belgian federal elections	Vlaams Belang (VB)

In addition, I created variables for other socio-economic indicators that might predict online support for ISIS. First, to examine whether local-level unemployment is linked to radicalization, I used official data on unemployment from France, Germany, the United Kingdom and Belgium, at the lowest possible level of aggregation. In France, Germany, and Belgium, the lowest possible level was the town/municipality. In the United Kingdom, data were available at the sub-municipality/neighborhood level.[22] I matched users to their respective areas for which unemployment data exist. As some have hypothesized that unemployment among *immigrants* in particular feeds ISIS radicalization (Grant, 2014), I also created a measure for the share of unemployed immigrants in each location. Panel B in Table 2.3 provides information on the distribution of these variables across Twitter users in the database.

Second, with the recent debates over the link between refugees and support for ISIS in Europe (Marans, 2015), I looked for variables that might proxy for the presence of refugees in a locality. I use information on the number of asylum seeker centers across localities in

[22]In the United Kingdom, statistical local-level data are available at the Mid-level Super Output Areas (MSOA), which are roughly the size of a neighborhood (Office for National Statistics, 2016).

France, and the share of asylum seeker benefits receivers in localities in Germany.[23] As these two variables are measured on different scales, I created a standardized measure for this combined variable. In addition, I use census data on the share of foreigners or non-citizens in each locality, to examine the extent to which ISIS supporters on Twitter are located in areas with higher shares of non-citizen populations. Table 2.3 shows the distribution of these variables across users. The Appendix provides more details on the data sources and construction of the independent variables.

2.4 Analysis

In this Section, I present the results of this study. First, in Subsection 2.4.1, I provide descriptive evidence that gives a sense for the nature of the data and its connection to real-world events. Next, in Subsection 2.4.2, I examine the cross-sectional link between far-right vote share and pro-ISIS radicalization. Finally, in Subsection 2.4.3, I study high-frequency changes in the wake of three events that stirred anti-Muslim sentiment and consider whether they increased online support for ISIS.

2.4.1 Descriptive, exploratory analysis

In this Subsection, I present a few examples that illustrate the kind of content that I collected and its connection to real-world events. On June 29, 2014, after conquering territories in Syria and Iraq, ISIS declared the establishment of a caliphate in an online statement distributed through Twitter and the group's media center,[24] calling all Muslims to pledge allegiance to it:

> *"So rush O Muslims and gather around your khalīfah, so that you may return as you once were for ages, kings of the earth and knights of war ... Come O*

[23]These data reflect 2014 figures.

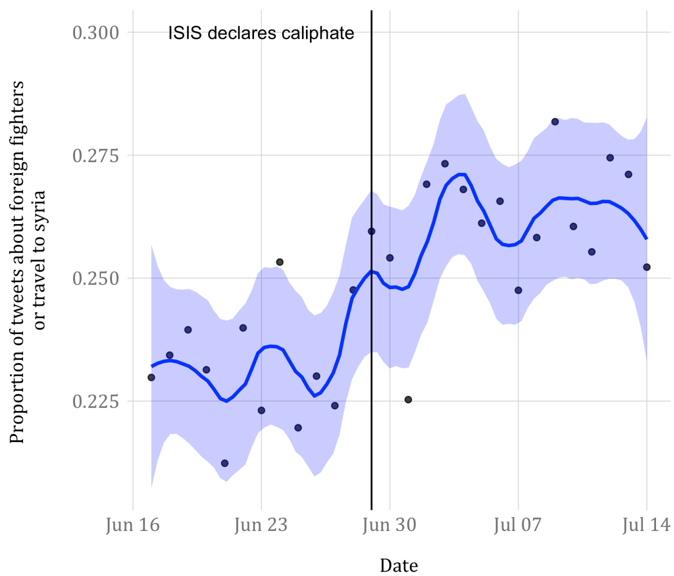
[24]Source: `http://myreader.toile-libre.org/uploads/My_53b039f00cb03.pdf`

> *Muslims to your honor, to your victory. By Allah, if you disbelieve in democracy, secularism, nationalism, as well as all the other garbage and ideas from the west, and rush to your religion and creed, then by Allah, you will own the earth, and the east and west will submit to you. This is the promise of Allah to you. This is the promise of Allah to you."*

I calculated the daily proportion of tweets discussing foreign fighters or travel to Syria posted by accounts located in France, Belgium, Germany, and the U.K. in the month surrounding ISIS's caliphate declaration. Figure 2.3 shows that after the declaration, discourse on foreign fighters significantly increased among these Twitter users.

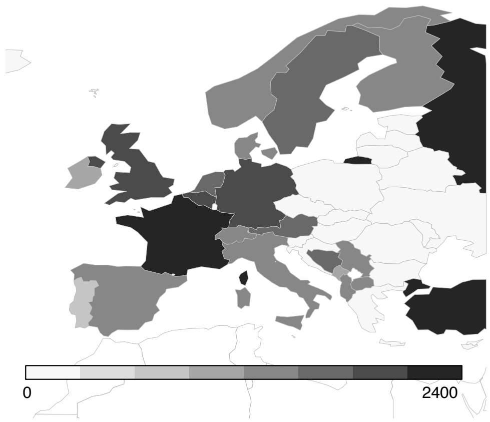
Next, I examine whether online radicalization measures correlate with Western foreign fighter figures. Figure 2.4 shows a map of ISIS foreign fighters from Europe (Panel a), along with a map showing the number of Twitter users flagged as ISIS activists by @CtrlSec in each country (Panel b). France, the United Kingdom, and Germany have higher numbers of foreign fighters and Twitters users flagged as ISIS activists than many other European countries. Figure 2.5 displays the correlation between additional online radicalization measures and the number of foreign fighters in the West. It can be seen that online measures of support for ISIS closely track official foreign fighter counts. While these scatterplots show bivariate relationships, the Appendix reports estimations controlling for population size, which show the same pattern.

Figure 2.3: ISIS declares caliphate and tweets discussing foreign fighters or travel to Syria

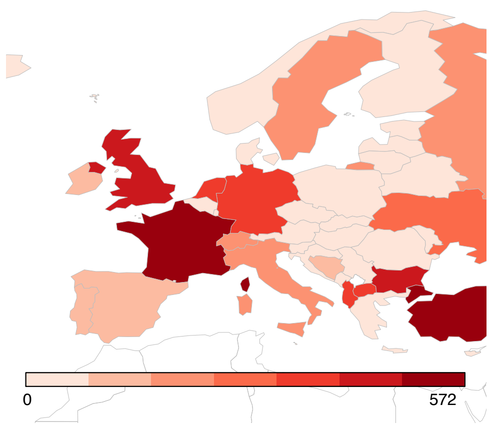


Note: The figure shows the daily proportion of tweets discussing foreign fighters or travel to Syria posted by accounts located in France, Belgium, Germany, and the U.K. in the month surrounding ISIS's caliphate declaration on June 29, 2014. The total number of tweets posted by these users during that month was 27,300, out of which 6,839 were labeled as discussing foreign fighters or travel to Syria.

Figure 2.4: Foreign fighters and online radicalization in Europe



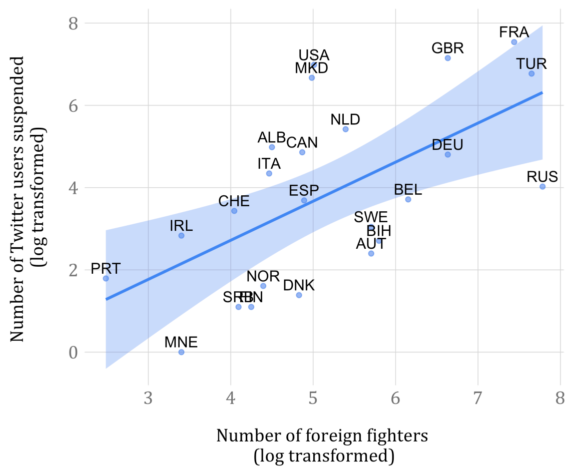
(a) Number of foreign fighters



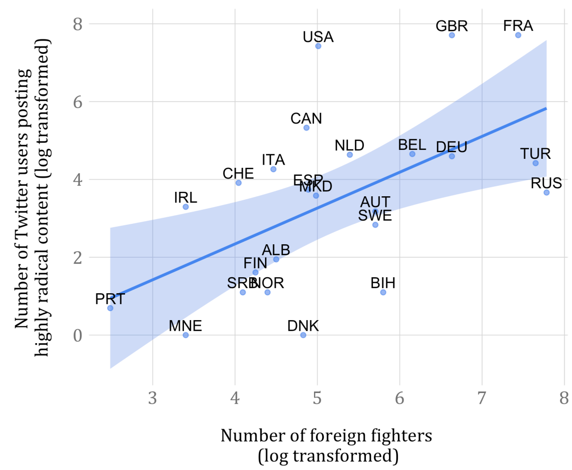
(b) Number of users flagged as ISIS activists

Note: Panel (a) displays official counts of ISIS foreign fighters in Europe, calculated by Barrett et al. (2015). Panel (b) shows the number of Twitter users flagged as ISIS activist by @CtrlSec, aggregated to the country level.

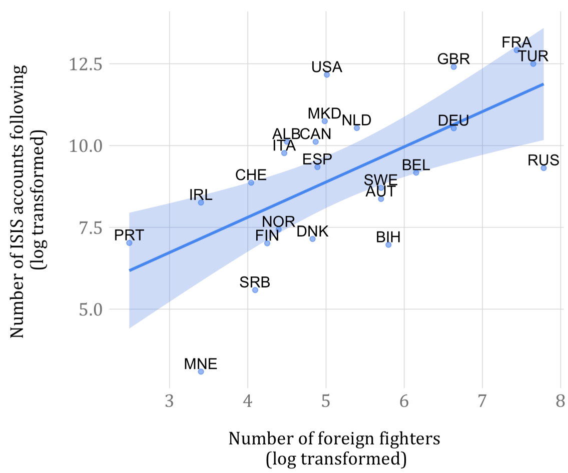
Figure 2.5: Foreign fighters and online radicalization (additional measures)



(a) Twitter suspension



(b) Highly radical content



(c) Number of ISIS accounts followed

Note: The figure plots scatterplots of the relationship between the number of foreign fighters and online radicalization measures in countries that had at least one foreign fighter with ISIS. Data on foreign fighters are taken from Barrett et al. (2015). Online radicalization measures are based on data collected by the author and are aggregated to the country level. The values are log-transformed.

2.4.2 Cross-sectional study: far-right vote share and pro-ISIS radicalization

In this Subsection, I examine whether local-level support for far-right parties is linked to greater pro-ISIS online radicalization. I regress the different online radicalization outcomes on the independent variables described in Subsection 2.3.4 using a combined dataset covering all localities in France, Germany, Belgium, and the United Kingdom. The dependent variables are summarized in Panel A in Table 2.3 and are measured on the Twitter user level. The independent variables, summarized in Panel B in Table 2.3, are matched to each individual user in the dataset, but originate in local-level administrative data.[25] I use the following least squares model in my main estimations:

$$Y_{ijk} = \beta_1 V_{ijk} + \beta_2 U_{ijk} + \beta_3 F_{ijk} + \beta_4 P_{ijk} + \beta_5 P_{ijk}^2 + \alpha_k C_k + \varepsilon_{jk} \tag{2.1}$$

Where i is a Twitter user in geographic area j in country k; Y_{ijk} is one of the online radicalization measures for user i in area j in country k; and V_{ijk} represents the locality-level vote share for far-right parties matched to user i in area j in country k. U_{ijk}, F_{ijk}, and P_{ijk} represent unemployment, share of foreigners, and population size matched to user i in area j in country k, respectively, and C_k is a country fixed effect.[26] The main coefficient of interest in these regressions is β_1, which estimates the relationship between the local-level vote share for far-right parties and online measures of support for ISIS. While this coefficient cannot be interpreted as evidence of a causal relationship, it provides a systematic test of

[25]To account for possible dependency across users in the same area, I cluster the standard errors at the locality level in my main regressions. Since the number of Twitter users in the database is much larger than the number of localities, clustering the standard errors drops 98% of the observations in the standard error calculation. Thus, I report in Section 2.6.7 in the Appendix results without clustered standard errors (but with heteroskedasticity robust standard errors). While these results should be taken with more caution, they are useful for considering relationships with noisy measures or when insufficient clusters, such as with country-by-country regressions.

[26]Data on the share of Muslim populations in each geographical area are only available in the United Kingdom. In estimations with United Kingdom data only, shown below, I find that controlling for Muslim population share does not affect the results.

the link between a context of anti-Muslim hostility and online pro-ISIS radicalization.

Far-right vote share and support for ISIS

Tables 2.5 and 2.6 report the main results. In Table 2.5, Column (1), the dependent variable is a text-based measure that is coded 1 for individuals who are at the top 1% of the distribution of posting pro-ISIS content, and 0 otherwise. To ensure that this content-based measure does not classify as 'radicalized' individuals who just post frequently on issues related to Islam, in Column (2) the dependent variable drops tweets about Islam, measuring 'radicalized content' only with tweets sympathizing with ISIS, describing life under ISIS territories, and discussing foreign fighters, the Syrian war, and anti-West sentiment. Regardless of the measure used, it can be seen that local-level vote share for far-right parties is positively associated with posting large numbers of radicalized tweets. In substantive terms, a one percent increase in right-wing vote share is associated with a 3% increase in the probability of being among the top 1% of posters of radical content.

Columns (3) - (5) in Table 2.5 report the results when the dependent variable is measured as being flagged by @CtrlSec as an ISIS activist, being suspended from Twitter for association with the organization, and with a count measure of the number of ISIS accounts that a user follows. Here, as well, the results show that vote share for far-right parties is positively related to these radicalization outcomes. However, suspension and number of ISIS accounts followed are not statistically significant at conventional levels with the clustered standard errors specification, although results are significant when estimating the models without clustered standard errors (see Table 2.29 in the Appendix). In substantive terms, vote share for far-right parties is associated with a 6% increase in the probability of being flagged as an ISIS activist.

Table 2.6 reports the results when the dependent variables reflect the number of tweets posted by a user across all six content outcomes. Here, a one percent increase in the vote share for far-right parties is positively and statistically significantly associated with increases in the number of tweets sympathizing with ISIS, relating to the life in ISIS-controlled territories, discussing the Syrian civil war, expressing anti-West sentiment, and reflecting faith in Islam. Substantively, these reflect an average increase of 4,000-9,000 pro-ISIS tweets across the entire sample. Note that these measures are calculated from content generated in English, Arabic, French, and German, and are measured across thousands of individuals in four countries. The consistency of the results across these text-based measures suggests that this association did not occur by random chance.

Other correlates of online radicalization

Next, I investigate other correlates of online radicalization. As can be seen in Tables 2.5 and 2.6, the unemployment rate at the local level is not robustly associated with online support for ISIS. While content-based outcomes are positively and significantly related to local-level unemployment (see Table 2.6), they are negatively related to unemployment when the dependent variable is measured as being flagged as an ISIS activist, being suspended from Twitter, or the number of ISIS accounts followed. To investigate whether local-level immigrant unemployment might drive online support for ISIS, I estimate regressions with a variable capturing the percent of unemployed immigrants in a locality in Table 2.7. Here, as well, results show that the share of unemployed immigrants is not significantly related to online radicalization.

In addition, I examine whether support for ISIS on Twitter relates to the share of foreigners or non-citizens in a locality. The third row in Tables 2.5 and 2.6 shows that a greater number of foreigners in a locality is positively associated with online radicalization, but the relationships are not statistically significant for almost all outcomes. In addition, I examine in Table 2.8 whether the share of refugees in a locality relates to greater support for ISIS on Twitter. I find that the share of asylum seeker and/or asylum seeker centers in a locality is negatively related to being flagged as an ISIS activist, being suspended from Twitter, and to the number of ISIS accounts followed. This is an important finding in light of recent debates over refugee policy in Europe, as it suggests that online radicalization on Twitter is not driven by the number of refugees in a locality.

To examine whether these results might be driven by a common third variable linked to both radicalization and far-right support, I use data on possible omitted variables that are available only in the U.K., such as the share of Muslims, Arabs, Pakistanis, Bangladeshis, and foreign-born in each local area. Table 2.9 shows that when controlling for these variables, vote share for far-right parties remains strongly correlated with the posting of radical pro-ISIS content on Twitter. The findings also show that the local proportion of Muslims is negatively and significantly correlated with posting pro-ISIS content. This is an important finding in light of recent debates on Muslim populations in the West, as it casts doubt on the argument that areas with larger Muslim populations are more likely to be prone to jihadi radicalization.

Overall, these results are consistent with the hypothesis that local-level vote share for far-right, anti-Muslim parties is linked to radicalization and support for the Islamic State on Twitter. The results hold across various dependent variables, in a large number of locations in four European countries. However, since the findings are based on cross-sectional comparisons, it is possible that these relationships are driven by reverse causality—i.e., that the presence of radicalized individuals in a locality increases support for far-right parties, and not the other way around. In the remaining parts of the paper, I investigate these relationships using high frequency Twitter data surrounding events that stir anti-Muslim sentiment: the Paris terrorist attacks in November 2015, the Brussels terrorist attacks in March 2016, and the PEGIDA anti-Muslim marches in February 2016.

Table 2.5: Socioeconomic correlates of support for ISIS on Twitter

	(1) Indicator for top 1% radical content	(2) Indicator for top 1% radical content, without Islam	(3) Flagged as an ISIS activist	(4) Suspended from Twitter	(5) Number of ISIS accounts following
Far-right vote share (%)	0.25**	0.25**	0.30**	0.09	86.48
	(0.12)	(0.12)	(0.14)	(0.39)	(82.92)
Unemployment (%)	0.20	0.18	-0.20	-1.24*	-111.71
	(0.25)	(0.25)	(0.52)	(0.69)	(140.11)
Foreigners (%)	0.10	0.11	0.26*	-0.06	84.01
	(0.10)	(0.09)	(0.15)	(0.32)	(69.89)
Constant	8.57*	7.68*	-9.76	35.07**	1116.19
	(4.56)	(4.42)	(6.34)	(15.24)	(3729.04)
Population controls	✓	✓	✓	✓	✓
Country fixed effects	✓	✓	✓	✓	✓
R^2	0.0003	0.0003	0.0062	0.0023	0.0062
Number of clusters	2,655	2,655	2,655	2,654	2,654
Number of observations	112,254	112,254	112,254	112,250	112,250

Robust standard errors in parentheses, clustered at the locality level. Base category is Belgium.
All coefficients are × 1,000 to account for the skewed distribution of the dependent variables.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.6: Socioeconomic correlates of posting pro-ISIS content on Twitter

	(1) Sympathy with ISIS	(2) Life in ISIS territories	(3) Travel to Syria or foreign fighters	(4) Syrian war	(5) Anti-West	(6) Islam
Far-right vote share (%)	0.05**	0.07**	0.07**	0.04**	0.04*	0.08*
	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.05)
Unemployment (%)	0.12**	0.14*	0.16**	0.11**	0.13**	0.23**
	(0.05)	(0.07)	(0.08)	(0.04)	(0.05)	(0.11)
Foreigners (%)	0.02	0.02	0.03	0.02	0.02	0.04
	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.04)
Constant	3.51***	5.53***	5.66***	2.68***	3.22***	7.49***
	(0.99)	(1.54)	(1.44)	(0.79)	(0.91)	(1.93)
Population controls	✓	✓	✓	✓	✓	✓
Country fixed effects	✓	✓	✓	✓	✓	✓
R^2	0.002	0.005	0.003	0.003	0.003	0.001
Number of clusters	2,655	2,655	2,655	2,655	2,655	2,655
Number of observations	112,254	112,254	112,254	112,254	112,254	112,254

Robust standard errors in parentheses, clustered at the locality level. Base country is Belgium.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.7: Unemployed immigrants and support for ISIS on Twitter

	(1) Indicator for top 1% radical content	(2) Indicator for top 1% radical content, without Islam	(3) Flagged as an ISIS activist	(4) Suspended from Twitter	(5) Number of ISIS accounts following
Unemployed immigrants (%)	0.09	0.11	0.30	-0.47	177.55
	(0.45)	(0.43)	(0.92)	(1.65)	(590.10)
Constant	14.89***	14.38***	0.43	26.64***	3990.51***
	(2.42)	(2.21)	(1.90)	(5.27)	(1327.63)
Population controls	✓	✓	✓	✓	✓
Country fixed effects	✓	✓	✓	✓	✓
R^2	0.0001	0.0001	0.003	0.001	0.001
Number of clusters	1,318	1,318	1,318	1,318	1,318
Number of observations	90,516	90,516	90,516	90,514	90,514

Robust standard errors in parentheses, clustered at the locality level. Base category is Belgium.
All coefficients are × 1,000 to account for the skewed distribution of the dependent variables.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.8: Asylum seekers and support for ISIS on Twitter

	(1) Indicator for top 1% radical content	(2) Indicator for top 1% radical content, without Islam	(3) Flagged as an ISIS activist	(4) Suspended from Twitter	(5) Number of ISIS accounts following
Asylum seekers (%, sd units)	0.12	0.07	-5.04**	-4.03	-675.54***
	(0.72)	(0.77)	(2.48)	(2.85)	(258.70)
Constant	5.96**	5.34**	14.12	50.98***	6852.84***
	(2.42)	(2.37)	(10.92)	(11.39)	(2384.48)
Controls	✓	✓	✓	✓	✓
Country fixed effects	✓	✓	✓	✓	✓
R^2	0.0001	0.0001	0.005	0.001	0.001
Number of clusters	1,209	1,209	1,209	1,209	1,209
Number of observations	88,388	88,388	88,388	88,386	88,386

Robust standard errors in parentheses, clustered at the locality level. Base category is Germany.
All coefficients are × 1,000 to account for the skewed distribution of the dependent variables.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.9: The local proportion of Muslims and the radical pro-ISIS posts in the U.K.

	(1) Sympathy with ISIS	(2) Life in ISIS territories	(3) Travel to Syria or foreign fighters	(4) Syrian war	(5) Anti-West
Far-right vote share (%)	0.06***	0.10***	0.09***	0.05***	0.04***
	(0.02)	(0.03)	(0.03)	(0.01)	(0.02)
Muslims (%)	-0.06**	-0.11**	-0.09**	-0.05**	-0.06**
	(0.03)	(0.05)	(0.04)	(0.02)	(0.02)
Males (%)	-0.04	-0.09	-0.10	-0.04	-0.05
	(0.04)	(0.07)	(0.07)	(0.03)	(0.04)
Pakistanis (%)	0.03	0.07*	0.06	0.03	0.03
	(0.02)	(0.04)	(0.04)	(0.02)	(0.02)
Bangladeshis (%)	0.00	0.04	0.01	0.02	0.01
	(0.02)	(0.04)	(0.04)	(0.02)	(0.02)
Arabs (%)	0.08	0.15	0.12	0.06	0.06
	(0.07)	(0.12)	(0.11)	(0.06)	(0.06)
Foreigners (%)	0.01	0.02	0.03*	0.01*	0.01*
	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
Unemployment (%)	-0.04	-0.07	-0.08	-0.03	-0.03
	(0.04)	(0.07)	(0.06)	(0.03)	(0.03)
Population	0.00***	0.00***	0.00***	0.00***	0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Population^2	-0.00***	-0.00***	-0.00***	-0.00***	-0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Constant	2.17	4.25	4.97	2.36	2.68
	(2.34)	(3.79)	(3.67)	(1.76)	(2.03)
R^2	0.001	0.001	0.001	0.001	0.001
Observations	61,925	61,925	61,925	61,925	61,925

Heteroskedasticity robust standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

2.4.3 Event studies: stirring anti-Muslim sentiment and support for ISIS

As discussed in Section 2.2, the relationship between anti-Muslim hostility and Islamist radicalization is complex, and likely runs in both directions. On one hand, anti-Muslim hostility might drive Muslim individuals to radicalize; on the other, radicalization, which in some cases results in violence and terrorism, can increase support for anti-Muslim, far-right parties. To further investigate how anti-Muslim hostility might drive radicalization and support for ISIS on social media, I take advantage of the high frequency nature of Twitter data and examine whether events that stir anti-Muslim sentiment increase online radicalization among potential ISIS recruits. This Subsection therefore adds to the prior analysis by studying changes in radicalized content over time.

In next subsections I describe these events. First, I consider whether the terrorist attacks in Paris (11/13/2015) and Brussels (3/22/2016) by ISIS-affiliated perpetrators affected radical pro-ISIS content in all localities in France, Germany, Belgium, and the United Kingdom. Second, I evaluate how an anti-Muslim event, the Patriotic Europeans Against the Islamization of the West (PEGIDA) movement's marches across Europe on February 6, 2016, may have led to greater online support for ISIS. As the two event studies reflect the same empirical design, I describe the estimations and results for both studies together.

Terrorist attacks

Terrorist attacks perpetrated by individuals associated with radical organizations are thought to increase anti-Muslim sentiment and violence in the West (Panagopoulos, 2006; Hanes and Machin, 2014). After 9/11, hate crimes against Muslims increased ten-fold in the United States (Gould and Klor, 2014). Anti-Muslim hostility also grew in Europe after recent terrorist attacks (Healy, 2015). The rise in anti-Muslim sentiment has benefited far-right parties in Europe: several polls conducted after the November 2015 attacks in Paris have shown that support for the far-right Front National party has substantially increased across France (Dearden, 2015; Todd, 2015). Following the Brussels attacks in March 2016, support on social media for the Belgian right-wing Vlaams Belang party increased 30-fold (Sykes, 2016).

I examine whether the Paris attacks of November 2015 and the Brussels attacks of March 2016 were immediately followed by increased radical, pro-ISIS content among Islamic State supporters on Twitter. The estimations and results are described below. One caveat with studying terrorist attacks, however, is that they might directly inspire radicalization. Individuals sympathetic to the Islamic State might feel inspired and emboldened by a 'successful' terrorist attack. This might lead to increased support for ISIS independent of anti-Muslim hostility. Thus, it is difficult to determine whether increases in radicalization following a terrorist attack are driven by anti-Muslim sentiment. For this reason, as described in the following section, I also examine an event that is likely to increase radicalization only through the anti-Muslim channel: the PEGIDA movement's marches across Europe in February 2016.

Paris terrorist attacks. On November 13, 2015, several perpetrators identified with the Islamic State launched several attacks in Paris, including suicide bombings and mass shootings. The attacks killed 130 people and injured hundreds of others, becoming the deadliest atrocities in France since the Second World War. Several polls conducted after the attacks have shown that support for the far-right Front National party has substantially increased across France (Dearden, 2015; Todd, 2015).

Brussels terrorist attacks. On March 22, 2016, ISIS-affiliated suicide bombers detonated explosive devices in Brussels Airport and at a train station nearby, killing 32 civilians and injuring over three hundred. After the attack, support on Facebook for the Belgian right-wing Vlaams Belang party increased by about three thousand percent (Sykes, 2016).

Anti-Muslim marches

The Patriotic Europeans Against the Islamization of the West (PEGIDA) movement was established in Dresden, Germany in October 2014 to oppose immigration, especially from Muslim-majority countries. On February 6, 2016, PEGIDA organized large marches in multiple cities in Germany, Britain, France, Netherlands, Austria, Ireland, Poland, Czech Republic, and Slovakia, to protest against the "Islamization of Europe" (Reuters, 2016). The marches drew thousands who came to express their opposition to the arrival of millions of migrants from Middle Eastern and North African countries, and to warn about Europe "being overrun by Muslims" (Reuters, 2016). This was the largest event organized by the movement so far (Meyer and Storck, 2015; *The Telegraph*, 2016).

The PEGIDA marches provide a useful event for testing the link between hostility against Muslims and pro-ISIS radicalization for two reasons. First, unlike terrorist attacks, which could have directly inspired radicalization, the anti-Muslim marches were likely to have affected radicalization only through the anti-Muslim channel. Second, since the marches did not provide any *new* evidence on the security threats of radicalized Muslims, they were unlikely to substantially change preexisting attitudes and increase anti-Muslim hostility in areas with low anti-Muslim sentiment. This contrasts with acts of terrorism, which have been shown to increase fear and anti-Muslim attitudes across the population as a whole (Huddy et al., 2005; Sides and Gross, 2013). In many places where the PEGIDA marches took place, counter-protesters gathered to oppose PEGIDA's positions (Huggler and Burgess, 2015; *RT News*, 2016). The marches can therefore be viewed as a "salience test"—an event that stirs emotions, but unlikely to dramatically change preexisting attitudes.[27]

[27] One caveat of studying the marches is that they were planned in advance, and were not a 'surprise' in time. Thus, it is important to not interpret the timing of the marches as exogenous. Nonetheless, the PEGIDA marches likely created an atmosphere where prior sentiment about Muslims became more salient, allowing the examination of variation across space in subsequent radicalization.

This implies different hypotheses about the effects of the terrorist attacks and the PEGIDA marches on radicalization across localities. First, the terrorist attacks should increase radicalization in general: both in areas that support far-right positions and areas that do not. On the other hand, the impact of the PEGIDA marches should be heterogeneous: radicalization should increase only in areas that already have high levels of anti-Muslim sentiment. In the following Subsections, I describe the estimation and results, respectively, for both of these studies.

Estimation

I estimate several difference-in-differences models to study both of these events, where I examine whether the difference in the number of pro-ISIS tweets 1-4 days after the event is larger in areas that have higher vote-share for far-right parties. To be sure, the terrorist attacks and the anti-Muslim marches were events that were discussed on national media to which everyone was likely exposed. The goal of this analysis is to examine whether changes in radical, pro-ISIS content after the events systematically varied between locations with low and high support for far-right parties. A 'pro-ISIS tweet' is coded 1 if its predicted value of belonging to any one of the six content categories—sympathy with ISIS, life in ISIS territories, travel to Syria or foreign fighters, Syrian war, anti-West, or Islam—is above the mean of the predicted values for that category, and 0 if not. For each event, I estimate the following least squares model:

$$Y_{ijk} = \beta_1 T_i + \beta_2 V_{ijk} + \beta_3 (T_i \times V_{ijk}) + \boldsymbol{\delta}\mathbf{X}_{ijk} + \alpha_k C_k + \varepsilon_{jk} \tag{2.2}$$

Where Y_{ijk} represents the level of radical content in tweet i posted in area j and country k, T_i is an indicator coded 1 for tweets appearing after the event (Paris attacks, Brussels attacks, and PEGIDA marches) and 0 if before, V_{ijk} is the locality-level vote share for far-right parties in area j in country k, $\mathbf{X}_{ijk}$ represents other independent variables described in equation (2.1), C_k represents country fixed effects, and ε_{jk} are standard errors clustered at the locality level.

Results

Tables 2.10 and 2.11 present the results for the Paris and Brussels terrorist attacks. In both tables, Column (1) uses data on tweets posted one day before and one day after the events; Column (2) uses data from two days before and after, and so forth. Panel A reports the change in the number of pro-ISIS tweets for the sample as a whole. It can be seen that in the first few days after the terrorist attacks in Paris and Brussels the number of radical tweets increased.

Panel B reports the results with the interaction between far-right vote share and the timing of the event. Results show that the difference between areas with low and high support for far-right parties is not statistically significant in most estimations. This pattern can be clearly seen in Figure 2.6, which plots the difference in the frequency of pro-ISIS tweets after the attack for areas with different levels of far-right vote share. The relatively flat line indicates that radicalized content increased in a similar manner across all localities in France, Germany, Belgium, and the United Kingdom, regardless of the levels of far-right vote share.

A different pattern can be seen for the anti-Muslim PEGIDA marches. Table 2.12 and Figure 2.7 show that the frequency of pro-ISIS tweets changed differently in areas with low and high support for far-right parties. While pro-ISIS content did not increase in areas with low far-right vote share after the PEGIDA marches, it significantly increased in areas with high far-right support. Figure 2.7 shows a sharp positive slope, where the frequency of pro-ISIS tweets significantly increased after the marches in areas with 30% far-right vote share or more. Figure 2.8 plots this pattern from data at the hourly level, showing the hourly proportion of pro-ISIS content produced by ISIS sympathizers in the three days before and after the PEGIDA marches. Here too, it can be seen that pro-ISIS content increased in areas with high levels of far-right support, but did not change in areas where far-right parties were not popular. While these results do not provide direct evidence that anti-Muslim hostility is responsible for the increase in pro-ISIS tweets in areas with high support for far-right parties, the results are consistent with the hypothesis that anti-Muslim animosity at the local level, expressed in support for far-right, anti-immigrant and anti-Muslim parties in Europe, increases online support for the Islamic State among potential sympathizers.

Table 2.10: The Paris attacks and pro-ISIS radicalization

	(1) [-1,+1]	(2) [-2,+2]	(3) [-3,+3]	(4) [-4,+4]
A. Changes in pro-ISIS radical content (standard deviation units)				
After attack = 1	0.160***	0.056***	0.044***	0.043***
	(0.022)	(0.011)	(0.011)	(0.008)
Far-right vote share (%)	-0.001	-0.001	-0.001	-0.000
	(0.003)	(0.001)	(0.001)	(0.001)
Constant	0.065	0.700***	0.741***	0.726***
	(0.148)	(0.058)	(0.058)	(0.052)
Controls	✓	✓	✓	✓
R^2	0.013	0.010	0.007	0.006
Number of clusters	268	327	362	386
Number of observations	9,150	15,223	21,459	27,637
B. Changes in radical content (standard deviation units), by far-right support				
After attack = 1	0.069**	0.086**	0.043	0.061**
	(0.031)	(0.036)	(0.035)	(0.027)
Far-right vote share (%)	-0.006*	-0.003	-0.004	-0.002
	(0.004)	(0.003)	(0.003)	(0.002)
After attack = 1 × Far-right vote share (%)	0.007***	0.002	0.004*	0.002
	(0.002)	(0.002)	(0.002)	(0.002)
Constant	0.160	0.107	0.213*	0.169
	(0.148)	(0.121)	(0.122)	(0.108)
Controls	✓	✓	✓	✓
R^2	0.014	0.011	0.007	0.007
Number of clusters	268	327	362	386
Number of observations	9,150	15,223	21,459	27,637

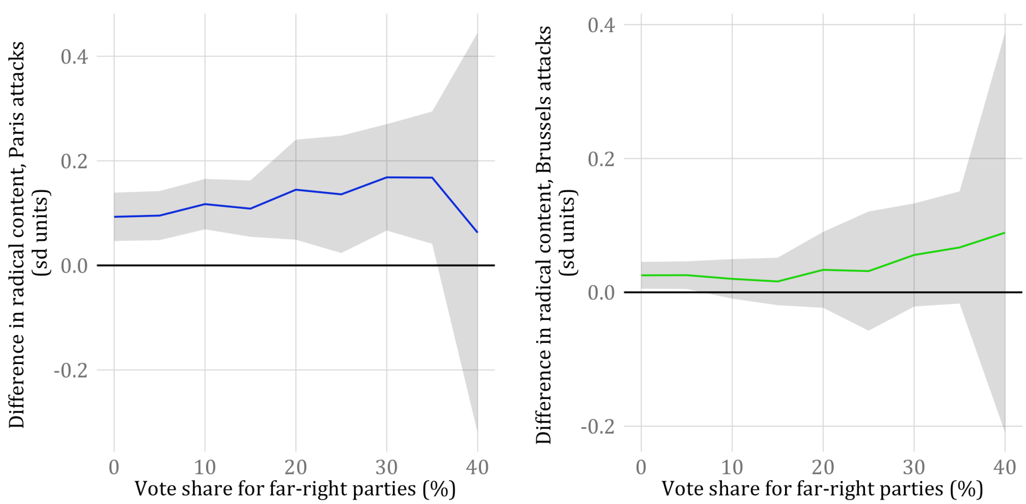
Note: Column (1) uses data on tweets posted one day before and one day after the events; column (2) uses data from two days before and after, and so forth.
Robust standard errors in parentheses, clustered at the locality level.
All coefficients are standardized.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.11: The Brussels attack and pro-ISIS radicalization

	(1) [-1,+1]	(2) [-2,+2]	(3) [-3,+3]	(4) [-4,+4]
A. Changes in pro-ISIS radical content (standard deviation units)				
After attack = 1	0.037***	0.049***	0.025**	0.013
	(0.014)	(0.013)	(0.010)	(0.009)
Far-right vote share (%)	0.002	0.002	0.001	0.002
	(0.002)	(0.002)	(0.002)	(0.002)
Constant	0.113	0.037	-0.001	0.019
	(0.102)	(0.082)	(0.076)	(0.077)
Controls	✓	✓	✓	✓
R^2	0.002	0.003	0.002	0.002
Number of clusters	392	472	529	571
Number of observations	17,613	32,164	46,460	60,773
B. Changes in radical content (standard deviation units), by far-right support				
After attack = 1	0.051**	0.058**	0.025	0.007
	(0.026)	(0.025)	(0.017)	(0.015)
Far-right vote share (%)	0.003	0.002	0.001	0.002
	(0.002)	(0.002)	(0.002)	(0.002)
After attack = 1 × Far-right vote share (%)	-0.001	-0.001	0.000	0.000
	(0.002)	(0.002)	(0.001)	(0.001)
Constant	0.103	0.029	-0.001	0.022
	(0.105)	(0.085)	(0.077)	(0.077)
Controls	✓	✓	✓	✓
R^2	0.002	0.003	0.002	0.002
Number of clusters	392	472	529	571
Number of observations	17,613	32,164	46,460	60,773

Note: Column (1) uses data on tweets posted one day before and one day after the events; column (2) uses data from two days before and after, and so forth.
Robust standard errors in parentheses, clustered at the locality level.
All coefficients are standardized.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure 2.6: The Paris and Brussels terrorist attacks and pro-ISIS radicalization, by vote-share for far-right parties



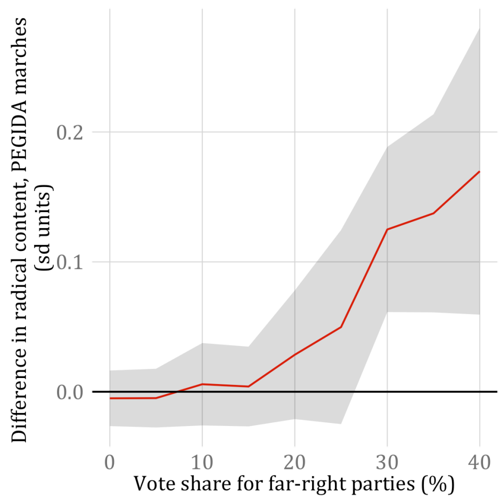
Note: The figure plots the difference in the frequency of pro-ISIS tweets after terrorist attacks for areas with different levels of far-right vote share. The differences are reported in standard deviation units. The left panel reports the results for the Paris attacks; the right panel shows the results for the Brussels attacks. Pro-ISIS content increased in a similar manner across all localities, regardless of the levels of far-right vote share.

Table 2.12: The PEGIDA marches and pro-ISIS radicalization

	(1) [-1,+1]	(2) [-2,+2]	(3) [-3,+3]	(4) [-4,+4]
A. Changes in radical content (standard deviation units)				
After PEGIDA marches = 1	-0.005	-0.012	-0.005	-0.009
	(0.020)	(0.012)	(0.011)	(0.008)
Far-right vote share (%)	-0.002	-0.000	0.000	0.001
	(0.002)	(0.002)	(0.002)	(0.001)
Constant	0.242**	0.090	0.135	0.143*
	(0.112)	(0.095)	(0.094)	(0.082)
Controls	✓	✓	✓	✓
R^2	0.003	0.002	0.002	0.002
Number of clusters	354	444	508	551
Number of observations	12,305	25,145	38,527	52,636
B. Changes in radical content (standard deviation units), by far-right support				
After PEGIDA marches = 1	-0.036	-0.044***	-0.038***	-0.028**
	(0.022)	(0.016)	(0.014)	(0.011)
Far-right vote share (%)	-0.003	-0.001	-0.001	-0.000
	(0.002)	(0.002)	(0.002)	(0.001)
After PEGIDA marches = 1 × Far-right vote share (%)	0.002	0.002*	0.002**	0.001*
	(0.002)	(0.001)	(0.001)	(0.001)
Constant	0.257**	0.103	0.149	0.151*
	(0.111)	(0.094)	(0.094)	(0.082)
Controls	✓	✓	✓	✓
R^2	0.003	0.002	0.002	0.002
Number of clusters	354	444	508	551
Number of observations	12,305	25,145	38,527	52,636

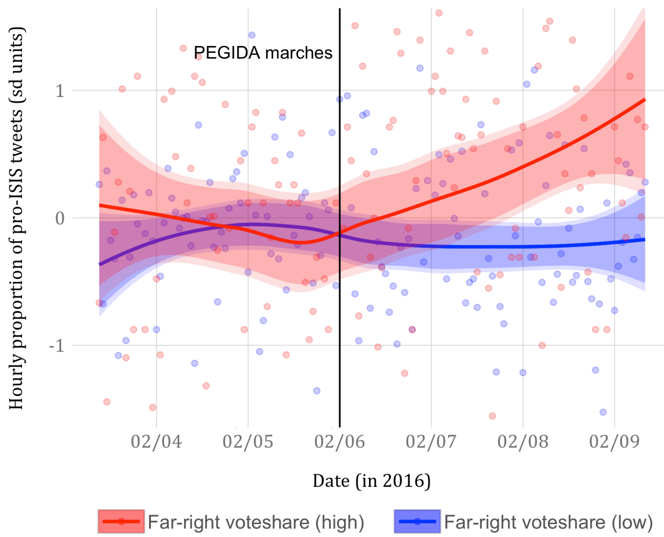
Note: Column (1) uses data on tweets posted one day before and one day after the events; column (2) uses data from two days before and after, and so forth.
Robust standard errors in parentheses, clustered at the locality level.
All coefficients are standardized.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure 2.7: The PEGIDA marches and pro-ISIS radicalization, by vote-share for far-right parties



Note: The figure plots the difference in the frequency of pro-ISIS tweets after the PEGIDA marches for areas with different levels of far-right vote share. The differences are reported in standard deviation units. Pro-ISIS content significantly increased after the marches in areas with 30% far-right vote share or more, but did not change in areas with low support for far-right parties.

Figure 2.8: The PEGIDA marches and pro-ISIS radicalization, by the hour



Note: The figure plots the proportion of pro-ISIS content produced by ISIS sympathizers in France, Germany, Belgium, and the United Kingdom (in standard deviation units) at the hourly level, in the three days before and after the PEGIDA marches. The red (blue) dots indicate content produced by individuals located in areas at the top (bottom) 10% of the voting for far-right parties. The bands mark 90% and 95% confidence intervals.

2.5 Conclusion

This study seeks to shed light on what drives so many to support the Islamic State in Europe. I argue that radicalization is part of a vicious cycle of reacting to, and feeding, anti-Muslim hostility. Socially-visible forms of radicalization, such as terrorist attacks, directly drive the vicious cycle, but other forms — like consuming ISIS content on social media — reflect a personal process whereby individuals are drawn to extremist ideology. By collecting data on thousands of Twitter accounts associated with ISIS, classifying millions of tweets along various dimensions of ISIS support, and mapping Twitter users to geographic locations in France, Germany, Belgium, and the United Kingdom, I showed that Twitter users located in areas that voted for far-right, anti-Muslim parties were more likely to show signs of radicalization than others in less hostile areas. While some have noted there might be a link between the rise of far-right parties and support for ISIS in Europe (Van Zeller, 2016), this paper has provided the first systematic, rigorous study of this proposition.

The findings stress the importance of understanding the process of radicalization and support for extremist movements in the age of social media. The ability to directly reach potential recruits on the Internet, interact with them through social media, and persuade them to embrace extremist ideology is changing how we think about recruitment in subnational conflicts. As the Internet and mobile technology continue to spread across the world, online radicalization is likely to continue, given the ongoing conflicts in the Middle East, North Africa, and other parts of the world. Studying how the online and offline worlds interact in this setting suggests that hostility in one's "offline" world might lead to the consumption of "online" radical content.

To be sure, this project leaves many questions unanswered, and in other parts of this dissertation, I examine additional aspects of the vicious cycle between radicalization and hostility. For example, in Chapter 3, I study the extent to which terrorist attacks perpetrated by Islamist groups increase anti-Muslim hostility and support for far-right parties in Europe and the United States. Of course, terrorist attacks are a highly visible form of radicalization, and thus play a key role in feeding the cycle of hostility and subsequent radicalization. But unlike the behavior studied in this project, locations targeted by terrorism are chosen to further strategic goals (Kydd and Walter, 2006). Indeed, these locations may not be home to many adherents of radical ideology, suggesting that terrorist violence might precipitate anti-Muslim hostility and subsequent radicalization in areas that were otherwise dormant.

Looking forward, research on radicalization would benefit from more localized studies aiming to causally identify the mechanisms by which anti-Muslim hostility is linked to online support for ISIS. Does an environment of anti-Muslim hostility increase online support for the Islamic State through a process of identity-seeking? Or is it driven by lack of opportunity to integrate into the surrounding society, e.g., by finding employment or increasing social status? Future work can also examine whether the patterns found in this paper are driven by individuals who are already ISIS sympathizers and, as a result of experiencing hostility, become more vocal in their support, or whether it is driven by the more moderate individuals who are pushed to radicalize after experiencing hostility. In addition, future work could study the determinants of ISIS radicalization in non-Western countries. While some of the same mechanisms might be at play, initial descriptive evidence suggests that recruits' motivations, as well as ISIS's recruitment strategy might be different in non-European countries (Wilson, 2015; Raghavan, 2016).

Finally, future studies might examine ways to de-radicalize potential recruits. With the rise of Islamic State recruitment on social media, several government agencies around the world have attempted to to counter ISIS messages online. While policy efforts such as the State Department's "Think Again Turn Away" campaign have had limited impact (Fernandez, 2015), other, more local and offline de-radicalization efforts have reportedly been more successful in this and prior conflicts (Rabasa et al., 2010; Horgan, 2015). In Chapter 4, I analyze over a hundred community engagement activities aimed to counter extremism in the United States, and evaluate whether they are associated with changes in pro-ISIS rhetoric by ISIS supporters in America. Better understanding of the process that leads individuals to sympathize with a foreign rebel group and radicalize could guide policymakers in identifying effective solutions to combat this troubling phenomenon.

2.6 Appendix

2.6.1 Identifying ISIS activist and follower accounts on Twitter

In this project, I track lists published publicly by several anti-ISIS hacking groups to identify ISIS supporters' accounts on Twitter. Using the Twitter APIs,[28] I designed an algorithm that continually monitors and records ISIS accounts identified by the hacktivist group @CtrlSec.[29] Immediately upon observing a new account in the @CtrlSec list, I download the complete "timeline" of tweets for the account, as well as its user profile, which includes various user-level fields, and list of the account's friends and followers. The full list of user profile fields is given in Table 2.13. The database contains "snapshots" of each user's profile at various points in time. In particular, prior to mid-May 2016, user profile snapshots were saved when the user was encountered on the @CtrlSec list or included as part of 5,000 randomly selected follower accounts for content sampling every 24 hours. Beginning in mid-May 2016, new snapshots are obtained for all non-suspended user accounts every 1-2 days, on average. The full list of data fields for each tweet is given in Table 2.14.

Downloading Twitter timelines

The dimensionality of the friends and followers is particularly challenging for historical timeline data collection. While I have identified approximately 15,000 activists thus far from the @CtrlSec postings, this has led to over 1.6 million followers and about 450,000 friends of these followers. Due to rate limits, it is impossible using the publicly available Twitter API to obtain full content timelines for 2 million accounts. Thus, I began by downloading the full historical tweet timelines of all @CtrlSec-identified "ISIS activist" accounts ($N = 14,979$), as well as of all the friends of a subsample of the activists who were first observed in the database as a follower or friend, and subsequently 'flipped' and

[28] `https://dev.twitter.com/overview/documentation`

[29] Lists are available in these handles: @ctrlsec, @ctrlsec0, @ctrlsec1, @ctrlsec2, @ctrlsec9.

became flagged as activists ($N = 193,973$). After completing an initial round of location prediction, I downloaded the complete historical tweet timelines of additional accounts of ISIS followers and friends predicted to be located in Europe and North America.

There are two additional sources of tweet timeline content in the dataset. The first is a so-called "random sample with holes." Since the Twitter Streaming API imposes rate limits on usage, I was only able to stream content for 5,000 users in a 24-hour period. The streaming began on December 19, 2015, and with the exception of occasional technical glitches, has been collecting data on the content posted by a random sample of 5,000 followers each day (data collection currently continues). Moreover, as noted previously, user profile information is downloaded at the same time. This ensures that user-level information (such as profile picture, number of friends, etc.), as well as account suspension status, are updated daily for this random sample.

The second source of tweet timeline data is a daily "total refresh" that began in May 2016. The Twitter API permits obtaining the a current profile snapshot for a user, which contains their most recently posted tweet, at a much faster rate limit than a full historical content download. Thus, I began to cycle through the entire database of nearly 2 million accounts on a daily basis, requesting latest profile and tweet, which leads to a complete refresh of user profiles and the latest tweet for each user in the system, as well as their suspension status, every 1-2 days on average. The total number of tweets scraped with this method was over 61 million as of August 2016.

Table 2.13: List of data fields at the user level

Field Name	Description
user_id	The integer representation of the unique identifier for this User.
date_added	The datetime the user profile snapshot was added to the database.
name	The name of the user, as they've defined it. Not necessarily a person's name.
screen_name	The screen name, handle, or alias that this user identifies themselves with.
location	The user-defined location for this account's profile. Not necessarily a location nor parseable.
description	The user-defined UTF-8 string describing their account.
url	A URL provided by the user in association with their profile.
protected	When true, indicates that this user has chosen to protect their Tweets.
followers_count	The number of followers this account currently has.
friends_count	The number of users this account is following (AKA their "followings").
listed_count	The number of public lists that this user is a member of.
created_at	The UTC datetime that the user account was created on Twitter.
favourites_count	The number of tweets this user has favorited in the account's lifetime.
utc_offset	The offset from GMT/UTC in seconds.
time_zone	A string describing the Time Zone this user declares themselves within.
geo_enabled	When true, indicates that the user has enabled the possibility of geo-tagging their Tweets.
verified	When true, indicates that the user has a verified account.
statuses_count	The number of tweets (including retweets) issued by the user.
lang	The BCP 47 code for the user's self-declared user interface language. May or may not have anything to do with the content of their Tweets.
profile_background_image_url	A HTTP-based URL pointing to the background image the user has uploaded for their profile.
profile_image_url	A HTTP-based URL pointing to the user's avatar image.
profile_image_file	A local copy of the user's profile image.
profile_banner_url	The HTTPS-based URL pointing to the standard web representation of the user's uploaded profile banner.
profile_banner_file	A local copy of the user's profile banner.
followers	The list of the user's followers, as of the date of this "snapshot." (Only obtained for certain users such as ISIS activists.)
friends	The list of the user's followers, as of the date of this "snapshot." (Only obtained for certain users such as ISIS activists.)
suspended	A flag for whether the account was suspended.

Note: Descriptions are copied verbatim from the Twitter REST API at https://dev.twitter.com/overview/api

Table 2.14: List of data fields at the tweet level

Field Name	**Description**
id	The integer representation of the unique identifier for this Tweet.
user_id	The integer representation of the unique identifier for the author of the Tweet.
date_added	The datetime that the Tweet was added to the database.
created_at	The datetime that the user account was created on Twitter.
text	The actual UTF-8 text of the status update.
source	Utility used to post the Tweet, as an HTML-formatted string. Tweets from the Twitter website have a source value of web.
truncated	Indicates whether the value of the text parameter was truncated, for example, as a result of a retweet exceeding the 140 character Tweet length. Truncated text will end in ellipsis, like this ...
in_reply_to_status_id	If the represented Tweet is a reply, this field will contain the integer representation of the original Tweet's ID.
in_reply_to_user_id	If the represented Tweet is a reply, this field will contain the integer representation of the original Tweet's author ID.
in_reply_to_screen_name	If the represented Tweet is a reply, this field will contain the screen name of the original Tweet's author.
retweet_count	Number of times this Tweet has been retweeted.
favorite_count	Indicates approximately how many times this Tweet has been "liked" by Twitter users.
lang	When present, indicates a BCP 47 language identifier corresponding to the machine-detected language of the Tweet text, or "und" if no language could be detected.
possibly_sensitive	This field is an indicator that the URL contained in the tweet may contain content or media identified as sensitive content.
coordinates	Represents the geographic location of this Tweet as reported by the user or client application.
withheld_in_countries	When present, indicates a list of uppercase two-letter country codes this content is withheld from.
quoted_status	This field only surfaces when the Tweet is a quote Tweet. This attribute contains the Tweet object of the original Tweet that was quoted.
retweeted_status	This attribute contains a representation of the original Tweet that was retweeted.

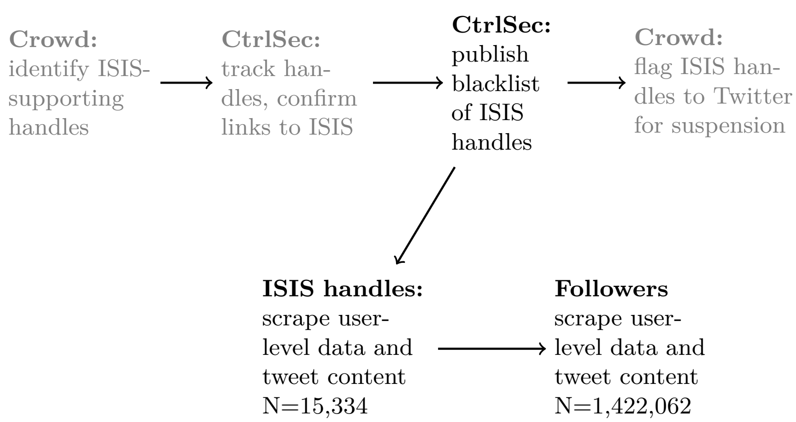
Note: Descriptions are copied verbatim from the Twitter REST API at https://dev.twitter.com/overview/api

Table 2.15: Number of tweets posted by all users in database, by year

Year	# tweets
2007	628
2008	3,389
2009	32,250
2010	84,913
2011	291,733
2012	826,552
2013	1,866,381
2014	3,980,438
2015	12,987,810
2016	47,107,004

Note: The number of tweets is accurate to 9/23/2016, 1:40PM ET.

Figure 2.9: Scraping ISIS accounts



Note: The number of Twitter users is accurate to 9/23/2016, 1:40PM ET

2.6.2 Predicting geographic location of ISIS activists and followers

Spatial Label Propagation algorithm

The spatial label propagation (SLP) algorithm used to predict the geographic locations of Twitter users in this paper implements the method developed by Jurgens (2013). The algorithm works as follows. First, define U to be a set of Twitter users in a social network, and for each user, let N be a mapping from the user to her friends (i.e., users to whom the user is directly connected), such that $u \rightarrow [n_i, ..., n_m]$. Also, let L be a mapping of users to their known geographic locations: $u \rightarrow (latitude, longitude)$, and E the current mapping from users to locations. E is being updated with each iteration of the algorithm.

The algorithm works as follows. First, it initializes E, the current mapping from users to locations, with L, the ground truth data. Then, for each user who does not have location data and has friends with location data, the algorithm creates a vector, M, which stores a list of the friends' locations. Using this list of latitude and longitude coordinates, the algorithm predicts the user's location by calculating the geometric median of the locations in M. The new predicted locations from the first round are added to E, the new mapping from users to locations. The algorithm repeats itself by predicting additional users' locations in the second round, using the ground truth and predicted location data from the previous round. The algorithm stops when the stopping criterion is met (in this paper, three rounds of prediction).

Figure 2.10 illustrates the way in which spatial label propagation algorithms work. First, location data from users who have them are used as "ground truth" to predict the locations of users to whom they are directly connected. If a user has more than one friend with ground truth data, the geometric median is calculated to predict his or her location. The geometric median is preferred over the geometric mean, as it represent the actual location of users in the network and not a meaningless average of coordinates. In addition, it is

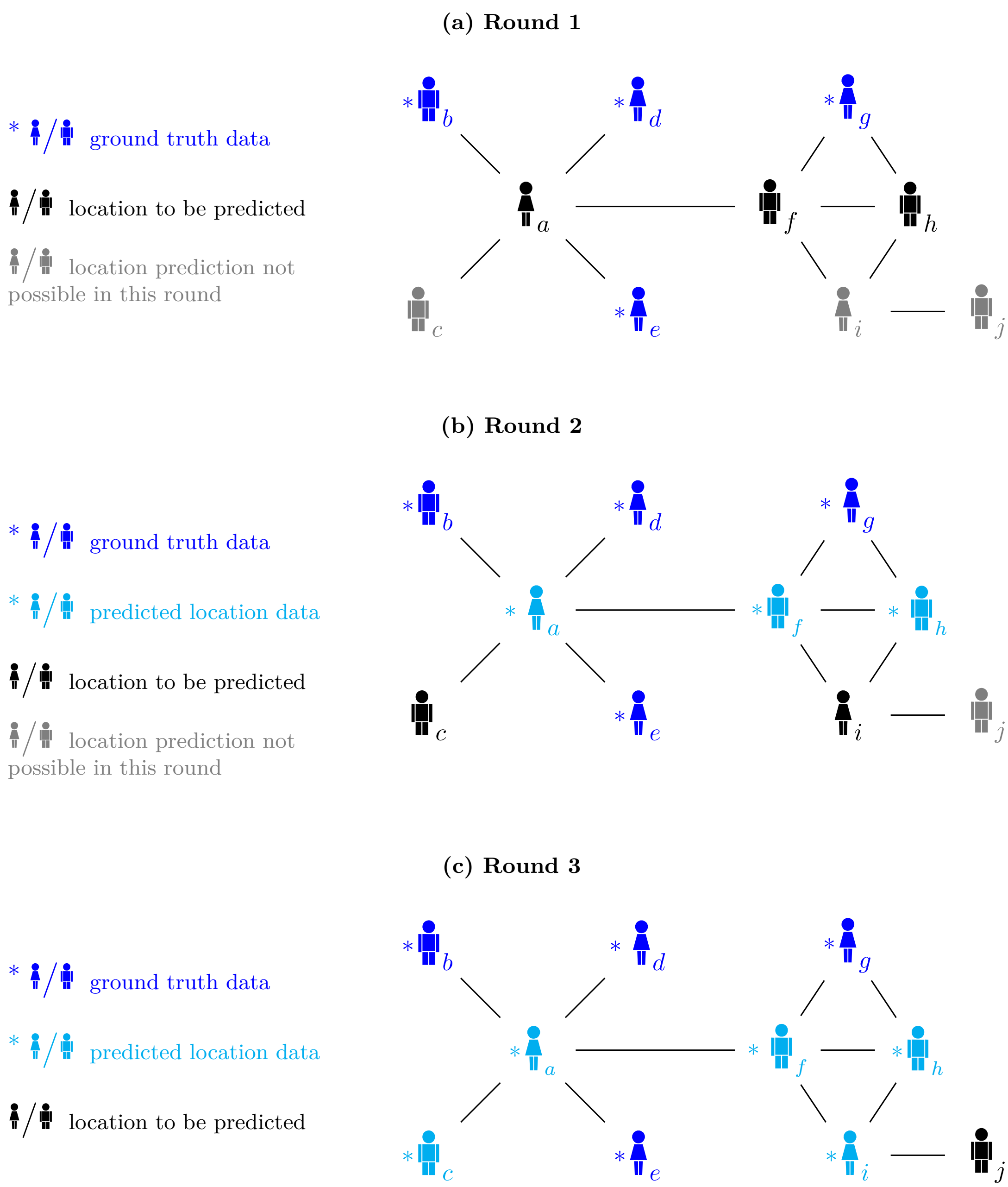
Data: U, L, and N
Let E be the current mapping from user to location;
Initialize E with L;
while *Convergence criteria are not met* **do**
    Let E' be the next mapping from user to (predicted) location;
    for $u \in (U - domain(L))$ *(i.e., users who do not currently have location information)* **do**
        Let M be a list of locations;
        for $n \in N(u)$ *(i.e., friends of user u)* **do**
            if $E(n) \neq \emptyset$ *(i.e., if the friend n has location information)* **then**
                add $E(n)$ to M;
            end
        end
        if $M \neq \emptyset$ *(i.e., user u's friends have location information)* **then**
            $E'(u) = \arg\min_{x \in L} \sum_{y \in L} distance(x, y)$ *(the predicted location of user u is the geometric median of her friends' locations)*
        end
    end
    $E = E'$
end
Result: Estimated user locations, E

Algorithm 1: Spatial Label Propagation (Jurgens, 2013)

less sensitive to outliers, which might happen when users post geo-located tweets while traveling. To give a concrete example, in Panel (a) the location of user a is predicted as the geometric median of users b, d, and e.

In the second stage, after the first round of prediction is completed and new users have predicted location information, the algorithm carries out a second round of location predictions, which uses richer location data that is distributed across the network, incorporating both ground truth and predicted location data points. Panel (b) shows that in the second round, it is possible to predict the location for user c using data on the location of users a, b, and e. In the same round, the location of user a is re-estimated, using a new data point from the predicted location of user f, in addition to the location information used in the first round, from users b, d, and e. This process is repeated a fixed number of times or until a minimum proportion of users have predicted location data.[30]

Figure 2.10: Spatial Label Propagation Algorithm



[30] I employed three iterations, which predicted locations for 1,626,350 users in the database.

I implement a slight deviation from the procedure described in Jurgens (2013). The original algorithm is designed to operate on a random sample of tweets, and not on a deep network of users who have timeline data and full lists of friends and followers. Thus, it identifies connections between individuals on the basis of "bidirectional mentions," i.e., user A mentions user B in a tweet and vice-versa. Bidirectional mentions are used in the original algorithm as a proxy for friends on social media, as it is impractical to obtain lists of friends and followers from a random sample of tweets. However, in my database, I have actual lists of friends and followers of accounts flagged as ISIS activists. As such, while I adopt the Jurgens (2013) algorithm as-is and allow connections between individuals to be identified on the basis of bidirectional mentions, I also generate "artificial" tweets containing bidirectional mentions between activists and their followers and friends. This ensures that the network structure contained in my database will be faithfully reproduced in the spatial label propogation algorithm.

The SLP algorithm requires so-called "ground truth" data, i.e., users with a known location, to base the prediction of the location for users without a known location. I obtained ground truth data as follows. For users with at least one geolocated tweet, I used the coordinates from an arbitrarily selected geolocated tweet. For users without any geolocated tweets but with a location field in their user profile, I looked up the location using the Google Maps and/or Bing Maps APIs (the specific API is selected arbitrarily).[31] If there was a match, I used the coordinates corresponding to this location as the user's ground truth location. To be sure, both of these methods are measured with error, but there is no reason to believe that these errors are systematically biased in any specific direction. Thus, by the law of large numbers, across the total universe of accounts with ground truth data ($N = 287,482$), these errors should be inconsequential.

[31] Google Maps API: `https://developers.google.com/places/web-service/details`; Bing Maps API: `https://msdn.microsoft.com/en-us/library/ff701711.aspx`.

Stability of location predictions

I verify the accuracy of the location prediction algorithm in the following way. The network structure in my database is relatively deep, centered around 14,979 ISIS activists for whom I have full lists of followers, as well as friends of a subset of the followers. Thus, individuals distributed across the network with ground truth data are connected to each other mainly through the ISIS activists' accounts. This is different from flat networks studied in other SLP applications using data from random samples of tweets (Jurgens et al., 2015). As a result, cross validation using only data from accounts with ground truth information is not useful for estimating the performance of the model.

In non-network data, cross validation on the training set is useful because observations do not depend on each other. Thus, $\hat{y_i}$, the prediction for observation i, is simply some function of the covariates for unit i and some parameters: $\hat{y_i} = f(x_i, \theta)$. Taking observations out in cross validation to test the model's prediction works well, because of the limited dependency between observations. In network data, cross validation is more problematic, because observations are dependent: $\hat{y_i} = f(\sum_j y_j, \theta)$. Therefore, taking observations out in cross validation does not only change θ, the parameters of the model, but also $\sum_j y_j$, the data used to predict $\hat{y_i}$. As a result, the estimations in the cross validation are likely to be biased, with greater bias for deeper networks in which the dependency between observations is higher.

To overcome this challenge and estimate the algorithm's performance, I designed a 10-fold out-of-sample stability test. I divided the training set into ten folds, and in each fold I randomly excluded 1/10 of the ground truth data when estimating the model. The algorithm therefore ran ten times, each time using only 90% of the training data to predict the locations of all users in the dataset ($N = 1,626,165$). I assume that the out-of-sample stability of the location prediction for each user i across ten folds can proxy the algorithm's location prediction accuracy. The logic behind this assumption is that highly unstable

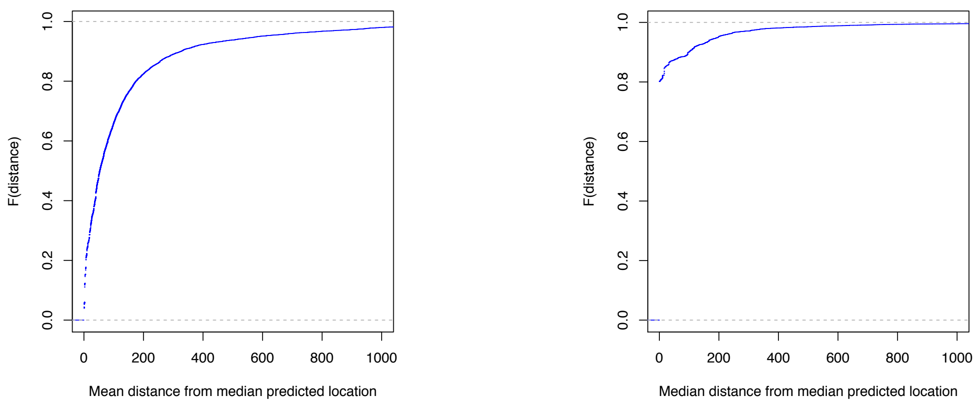
(stable) predictions across ten different prediction exercises likely means that the prediction is not very accurate (accurate). If a given user's friends are distributed geographically in a manner that renders the prediction highly unstable when excluding a random portion of the friends, then it means that the geometric median of the friends' locations is probably not a good proxy for the user's true location. On the other hand, if leaving out friends with location data does not affect the stability of the user's predicted location, then it means that many of the user's friends are located in the same area, making prediction stable, and likely more accurate.

After obtaining ten different location predictions for each user in the dataset, I calculated, for each user i, the mean and median distance from the median location predicted for user i. Figure 2.11 shows the performance for the ISIS activists' accounts ($N = 14,979$). Figure 2.12 shows the performance for the ISIS followers' accounts ($N = 1,611,633$). The figures plot the cumulative distribution function of the location predictions' stability across ten prediction estimations. In Panel (a), the stability is calculated as the mean of the predicted locations' deviations from the median predicted location for each user across the ten folds. In Panel (b), the stability is calculated as the median of the predicted locations' deviations from the median prediction. When using the mean stability measure, the majority of users' predicted locations are stable around a radius of about 50 kilometers or less for activists, and 70 kilometers or less for followers. When using the median stability measure, for over 80% of the users locations are predicted with a median stability of 10 kilometers or less.

Prediction stability and the study's covariates

To examine the extent to which prediction stability might lead to systematic bias in the study's point estimates, I regressed the mean and median prediction stability measures on the study's covariates, replicating model (1) in the main paper. The results are reported

Figure 2.11: 10-Fold out-of-sample stability test (ISIS activists' accounts)

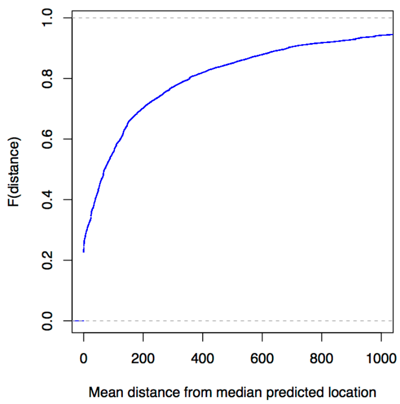


(a) Mean stability (b) Median stability

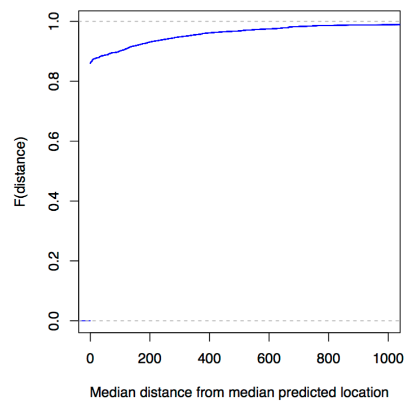
Note: The figure plots the cumulative distribution function of the stability of location predictions of ISIS activists ($N = 14,979$) across ten prediction estimations when leaving out one-tenth of the training data each time. In Panel (a), the stability is calculated as the mean of the predicted locations' deviations from the median predicted location for each user across the ten folds. The x axis shows the mean distance from the median predicted location for each user. The y axis shows the probability that mean deviation is x distance or less from the user's median predicted location. In Panel (b), the stability is calculated as the median of the predicted locations' deviations from the median prediction. When using the mean stability measure, the majority of users' predicted locations are stable around a radius of about 50 kilometers or less. When using the median stability measure, for over 80% of the users locations are predicted with a median stability of 10 kilometers or less.

in Table 2.16. Local-level vote share for far-right parties does not systematically correlate with any of the prediction stability measures. Other covariates, such as population density and country dummies also do not correlate with prediction stability. However, the percent of unemployment and the percent of foreigners negatively correlate with the prediction stability, meaning that accounts predicted to be located in areas with higher unemployment and share of foreigners tend to have less stable predictions. This means that the results reported in the main paper for these covariates should be taken with more caution, and that controlling for these variables is important for accounting for this prediction instability.

Figure 2.12: 10-Fold out-of-sample stability test (ISIS followers' accounts)



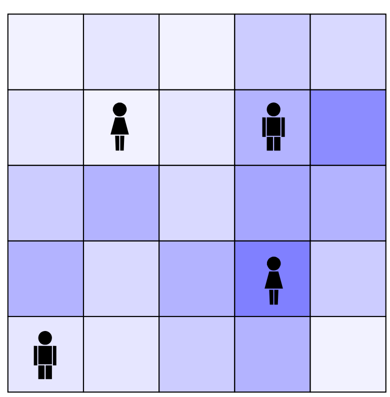
(a) Mean stability



(b) Media stability

Note: The figure plots the cumulative distribution function of the stability of location predictions of ISIS followers ($N = 1,611,633$) across ten prediction estimations when leaving out one-tenth of the training data each time. In Panel (a), the stability is calculated as the mean of the predicted locations' deviations from the median predicted location for each user across the ten folds. The x axis shows the mean distance from the median predicted location for each user. The y axis shows the probability that mean deviation is x distance or less from the user's median predicted location. In Panel (b), the stability is calculated as the median of the predicted locations' deviations from the median prediction. When using the mean stability measure, the majority of users' predicted locations are stable around a radius of about 70 kilometers or less. When using the median stability measure, for over 80% of the users locations are predicted with a median stability of 10 kilometers or less.

Figure 2.13: Matching Twitter users to electoral constituencies



Note: The figure illustrates how I used predicted geo-location data to match users to locations with various levels of far-right vote share. In the figure, darker shades reflect areas with higher far-right vote share.

Table 2.16: 10-Fold out-of-sample stability and the study's covariates

	(1) Mean stability (km)	(2) Median stability (km)
Vote share for far-right parties (%)	-0.35	0.44
	(2.90)	(0.97)
Unemployment (%)	-16.92**	-3.25*
	(6.64)	(1.90)
Foreigners (%)	-4.76*	-1.53*
	(2.52)	(0.90)
Population	0.00	0.00
	(0.00)	(0.00)
Population^2	-0.00	-0.00
	(0.00)	(0.00)
Germany dummy	-46.16	-33.07
	(116.35)	(40.96)
France dummy	-212.74*	-70.67
	(115.04)	(44.46)
U.K. dummy	-77.65	-20.06
	(93.90)	(28.19)
Constant	551.68***	135.89***
	(111.70)	(38.84)
R^2	0.027	0.005
Number of clusters	2,652	2,652
Number of observations	112,229	112,229

Robust standard errors in parentheses, clustered at the locality level. Base country is Belgium.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

2.6.3 Classifying Twitter content

To generate the textual content outcomes in this study, I used supervised machine learning to classify tweets into several categories: (1) Anti-West, (2) Islam, (3) Sympathy with ISIS, (4) Life in ISIS territories, (5) Travel to Syria or foreign fighters, and (6) Syrian war. For each of the four languages: English, Arabic, French and German, I obtained a random sample of tweets posted by ISIS activists (i.e., the accounts that have been flagged by @CtrlSec). These tweets served as a training set for a classification model. The sizes of the training sets varied by language: English ($N = 9,926$), Arabic ($N = 10,631$), French ($N = 6,158$), and German ($N = 3,011$). Each tweet was assigned one or more of the categories by three distinct Amazon Mechanical Turk and/or Crowdflower workers, and label(s) were retained for a given tweet if and only if there was "majority agreement," i.e., at least two out of the three workers assigned the same label(s) to the tweet.

After obtaining the training set labels, I pre-processed the tweet text as follows. For tweets in the English, French and German languages, I removed punctuation, numbers, stop words, and applied standard word stemming algorithms for each language. For tweets in the Arabic language, I similarly removed punctuation and numbers. To pre-process Arabic tweets, I applied a standard set of Arabic text preparation techniques.[32]

With the pre-processed text, I generated a document-term matrix composed of unigrams and bigram tokens. That is, I obtained the frequency of individual words and two-word phrases that appeared in these tweets. I combined unigrams and bigrams in order to provide more textual structure and increase the predictive accuracy of the models. Any term included in the document-term matrix must have had appeared in at least two tweets

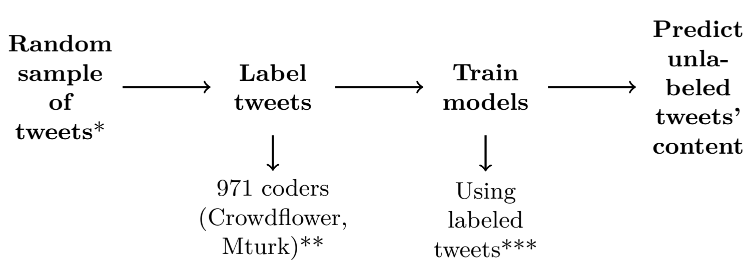
[32]Specifically, I removed leading 'alif lam' with optional leading 'waw'; leading 'alif lam' or double 'lam' at start of the text; leading 'kaf alif lam' with optional 'waw'; leading 'ba alif lam' with optional 'waw'; leading 'fa alif lam' with optional 'waw'; leading double 'alif' with optional 'lam' and an optional leading 'waw'; trailing 'ha,' 'ya ya nun,' 'ya waw nun,' 'ha' or 'ha alif,' 'ha mim,' 'ha mim alif'; and single letters such as 'waw.' I used the code from: `http://badhessian.org/2012/08/text-normalization-and-arabic-in-r/`

in order to be included in the classification model. Then, I applied a term-frequency / inverse-document-frequency (tf-df) transformation to down-weight the frequency of very common phrases across the whole corpus, as is standard in automated content analysis (Ramos, 2003).

Since Twitter textual data are very noisy, and radical pro-ISIS content is rare, many tweets in the database were coded as unrelated to any of the above categories. Class proportions for each language in the training set are shown in Tables 2.17 – 2.20. To facilitate statistical prediction, I followed King and Zeng (2001), randomly over-sampling pro-ISIS tweets and randomly under-sampling unrelated tweets to obtain a class proportion of 0.5 for each of the categories, for each topic, for each language.

I trained separate logit models using the labeled rebalanced training sets for each category in each language. For all specifications, I used the the elastic-net generalized linear model (Friedman, Hastie and Tibshirani, 2010), selecting the regularization parameter λ by cross-validation to maximize the area under the ROC curve. Figures 2.15 – 2.18 show the cross-validation curves for each language and topic. Model performance statistics for each topic and language are shown in Tables 2.21 – 2.20. The classification models for each topic and language were then employed on the full set of tweets in the database to classify each unlabeled tweet as belonging to one or more of these categories.

Figure 2.14: Supervised machine learning



Note: * English: 9,926; Arabic: 10,631; French: 6,158; German: 3,011.
** Each tweet coded by 3 coders, label retained if there was majority agreement.
*** Over-sample pro-ISIS content, under-sample unrelated tweets.

Table 2.17: Class proportions by topic (English)

	0	1
Anti-West	0.984577	0.015423
Islam	0.858215	0.141785
Sympathy with ISIS	0.982727	0.017273
Life in ISIS territories	0.963603	0.036397
Travel to Syria or foreign fighters	0.996607	0.003393
Syrian war	0.924532	0.075468

Table 2.18: Class proportions by topic (Arabic)

	0	1
Anti-West	0.998104	0.001896
Islam	0.913460	0.086540
Sympathy with ISIS	0.996777	0.003223
Life in ISIS territories	0.996777	0.003223
Travel to Syria or foreign fighters	0.999526	0.000474
Syrian war	0.981043	0.018957

Table 2.19: Class proportions by topic (French)

	0	1
Anti-West	0.971370	0.028630
Islam	0.890500	0.109500
Sympathy with ISIS	0.965607	0.034393
Life in ISIS territories	0.965607	0.034393
Travel to Syria or foreign fighters	0.982711	0.017289
Syrian war	0.947388	0.052612

Table 2.20: Class proportions by topic (German)

	0	1
Anti-West	0.959585	0.040415
Islam	0.924352	0.075648
Sympathy with ISIS	0.932124	0.067876
Life in ISIS territories	0.915026	0.084974
Travel to Syria or foreign fighters	0.947668	0.052332
Syrian war	0.915026	0.084974

Table 2.21: Model performance (English)

	anti-west	islamic-faith	is-sympathy	is-life	syria-travel-ff	syrian-war
Sensitivity	0.9981	0.9363	0.9992	0.9971	0.9992	0.9441
Specificity	1.0000	0.6752	0.9948	0.9948	1.0000	0.9943
Pos Pred Value	1.0000	0.7498	0.9949	0.9949	1.0000	0.9939
Neg Pred Value	0.9982	0.9106	0.9992	0.9971	0.9992	0.9478
Prevalence	0.4962	0.5097	0.5064	0.5024	0.5086	0.4949
Detection Rate	0.4953	0.4772	0.5060	0.5009	0.5082	0.4672
Detection Prevalence	0.4953	0.6364	0.5085	0.5035	0.5082	0.4701
Balanced Accuracy	0.9991	0.8057	0.9970	0.9960	0.9996	0.9692

Table 2.22: Model performance (Arabic)

	anti-west	islamic-faith	is-sympathy	is-life	syria-travel-ff	syrian-war
Sensitivity	0.9985	0.9627	0.9985	0.9987	0.9991	0.9583
Specificity	1.0000	0.9924	1.0000	1.0000	1.0000	1.0000
Pos Pred Value	1.0000	0.9922	1.0000	1.0000	1.0000	1.0000
Neg Pred Value	0.9985	0.9634	0.9985	0.9987	0.9990	0.9599
Prevalence	0.5039	0.5028	0.5094	0.4973	0.5115	0.5007
Detection Rate	0.5031	0.4841	0.5086	0.4967	0.5110	0.4798
Detection Prevalence	0.5031	0.4879	0.5086	0.4967	0.5110	0.4798
Balanced Accuracy	0.9992	0.9775	0.9993	0.9993	0.9995	0.9792

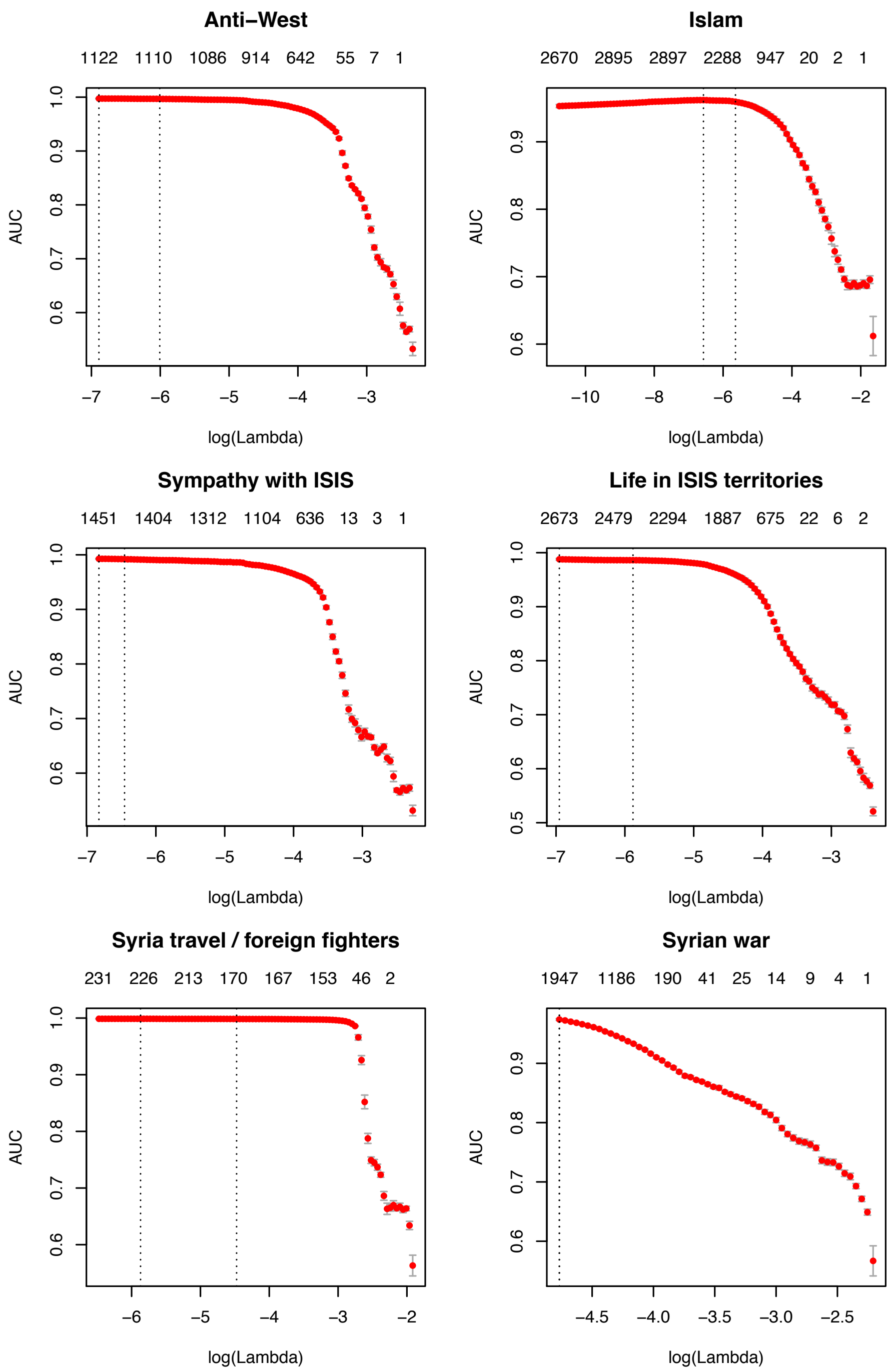
Table 2.23: Model performance (French)

	anti-west	islamic-faith	is-sympathy	is-life	syria-travel-ff	syrian-war
Sensitivity	0.9985	0.9910	0.9925	0.9876	0.9985	0.9926
Specificity	0.9923	0.9978	0.9952	1.0000	1.0000	1.0000
Pos Pred Value	0.9922	0.9977	0.9951	1.0000	1.0000	1.0000
Neg Pred Value	0.9985	0.9912	0.9926	0.9872	0.9985	0.9926
Prevalence	0.4951	0.4980	0.4975	0.5114	0.5127	0.5031
Detection Rate	0.4943	0.4936	0.4938	0.5051	0.5120	0.4993
Detection Prevalence	0.4982	0.4947	0.4962	0.5051	0.5120	0.4993
Balanced Accuracy	0.9954	0.9944	0.9939	0.9938	0.9993	0.9963

Table 2.24: Model performance (German)

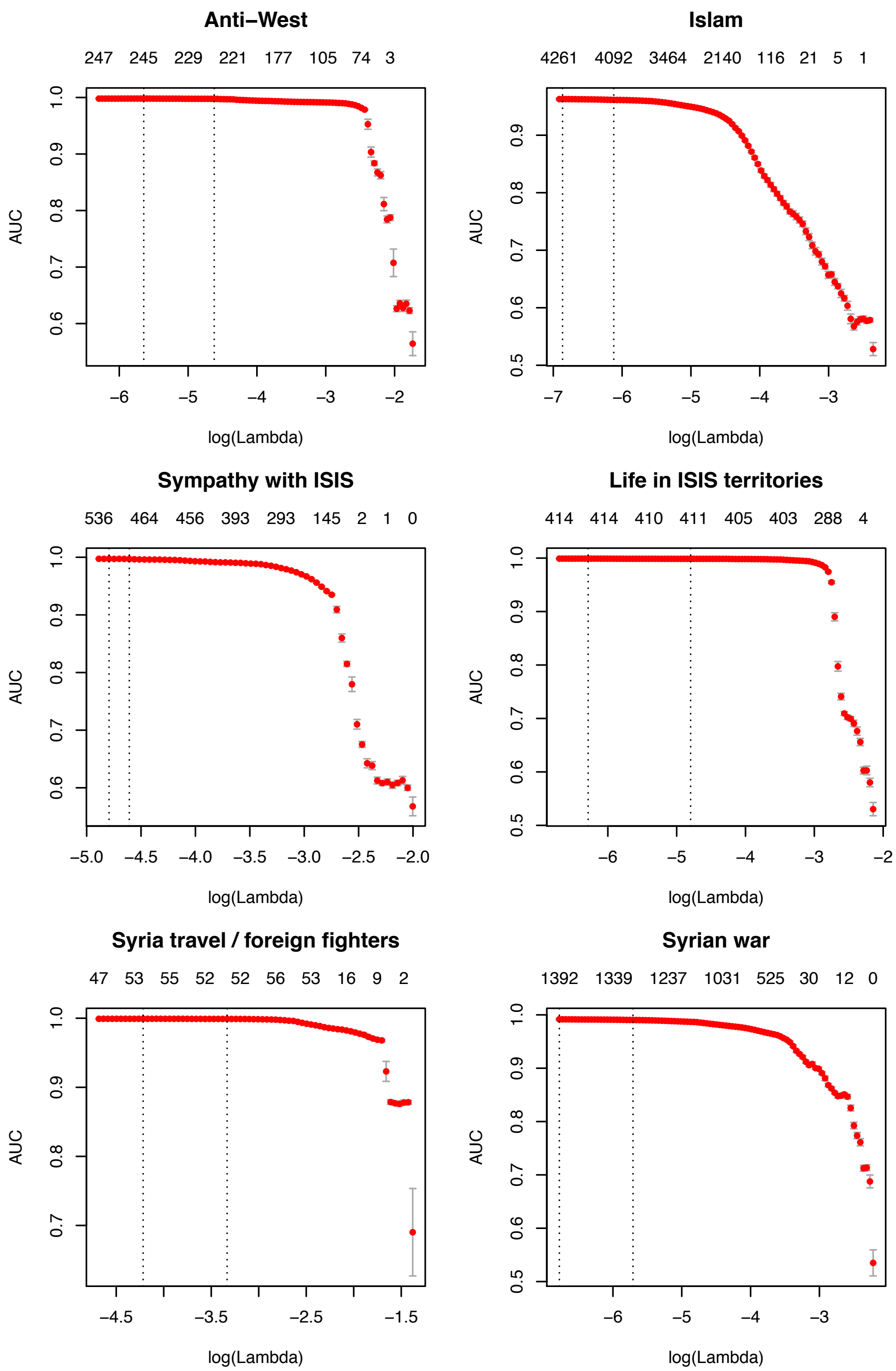
	anti-west	islamic-faith	is-sympathy	is-life	syria-travel-ff	syrian-war
Sensitivity	0.9661	0.9828	0.9839	0.9770	0.9818	0.9720
Specificity	1.0000	0.9713	0.9766	0.9949	0.9766	0.9880
Pos Pred Value	1.0000	0.9729	0.9779	0.9947	0.9778	0.9869
Neg Pred Value	0.9686	0.9818	0.9829	0.9778	0.9808	0.9744
Prevalence	0.4891	0.5119	0.5135	0.4964	0.5124	0.4813
Detection Rate	0.4725	0.5031	0.5052	0.4850	0.5031	0.4679
Detection Prevalence	0.4725	0.5171	0.5166	0.4876	0.5145	0.4741
Balanced Accuracy	0.9831	0.9771	0.9802	0.9859	0.9792	0.9800

Figure 2.15: Cross validation for model choice (English tweets)



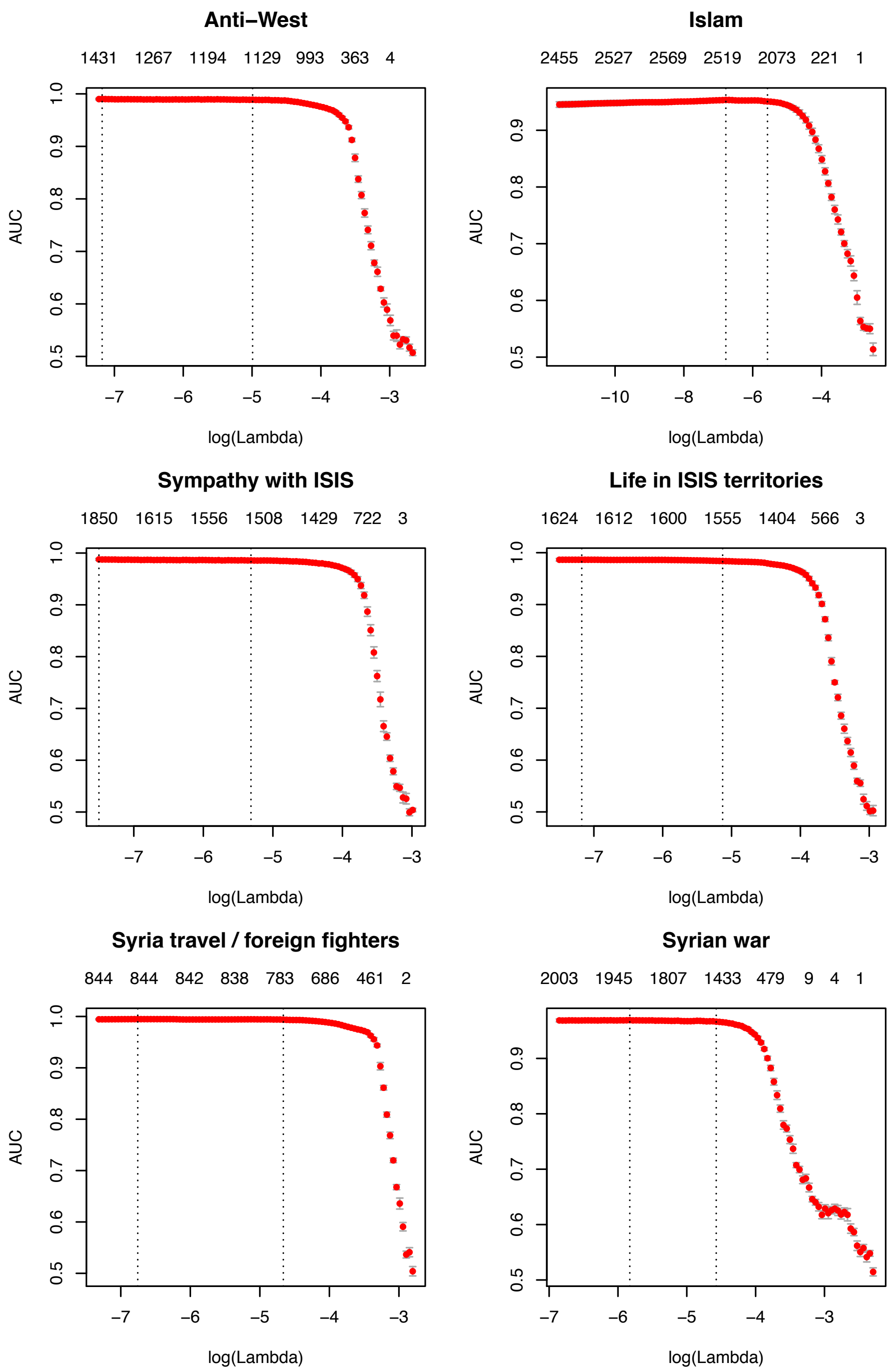
Note: The figure shows cross-validation curves for model choice in text classification of English language tweets for six topics. The cross-validation estimates for each model are shown in red dots, surrounded by error bars, plotted against the λ sequence. The y axis marks the Area Under the ROC Curve (AUC). Two selected λs are marked by vertical dotted lines. The numbers at the top of the figures represent the number of tokens (unigrams and bigrams) used in each model.

Figure 2.16: Cross validation for model choice (Arabic tweets)



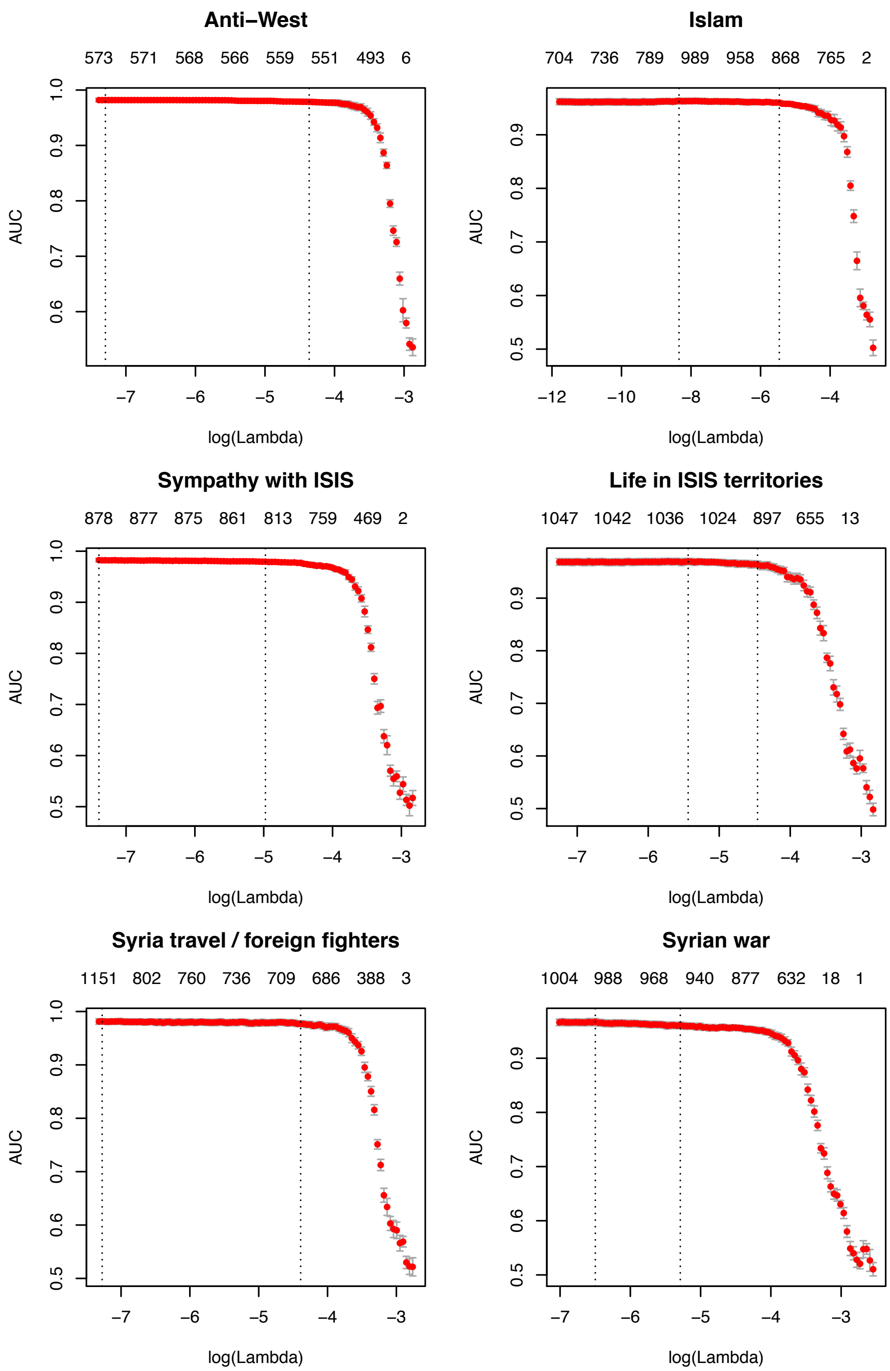
Note: The figure shows cross-validation curves for model choice in text classification of Arabic language tweets for six topics. The cross-validation estimates are shown in red dots, surrounded by error bars, plotted against the λ sequence. The y axis marks the Area Under the ROC Curve (AUC). Two selected λs are marked by vertical dotted lines. The numbers at the top of the figures represent the number of tokens (unigrams and bigrams) used in each model.

Figure 2.17: Cross validation for model choice (French tweets)



Note: The figure shows cross-validation curves for model choice in text classification of French language tweets for six topics. The cross-validation estimates are shown in red dots, surrounded by error bars, plotted against the λ sequence. The y axis marks the Area Under the ROC Curve (AUC). Two selected λs are marked by vertical dotted lines. The numbers at the top of the figures represent the number of tokens (unigrams and bigrams) used in each model.

Figure 2.18: Cross validation for model choice (German tweets)



Note: The figure shows cross-validation curves for model choice in text classification of German language tweets for six topics. The cross-validation estimates are shown in red dots, surrounded by error bars, plotted against the λ sequence. The y axis marks the Area Under the ROC Curve (AUC). Two selected λs are marked by vertical dotted lines. The numbers at the top of the figures represent the number of tokens (unigrams and bigrams) used in each model.

2.6.4 Collecting administrative data from European countries

To assign independent variables to each user in my database, I collected administrative data from France, Germany, Belgium and the United Kingdom on far-right vote share, percent unemployment, share of foreigners, population size, and additional variables described below. I matched each variable to its corresponding spatial polygon using shape files from official government databases. Then, I used Twitter users' predicted geo-location data and the shape files of local administrative areas to assign users to areas with local-level socioeconomic data. This process was done in `R`, and the code to replicate the point-to-polygon matching is available upon request.

2.6.5 Far-right vote share

France I obtained data on voting results in the 2015 French Departmental Elections at the polling station level from France's open platform of public data.[33]. The data contain information on the votes for each party in each polling station, the total eligible votes, as well as the electoral canton in which each polling station is located, among other variables. I aggregated the votes for the Front National party to the electoral canton level, and then divided the raw vote total for the party by the total eligible votes in each electoral canton. I used the electoral canton level vote share because of the availability of shape files at that level.

Germany I obtained data on voting results in the 2013 Federal Elections in Germany at the constituency level from Germany's Federal Returning Officer's Office.[34] For each constituency, I calculated the percent vote share in the Second Vote for the National Democratic Party of Germany (NPD) and the Alternative for Germany (AfD) party.

---

[33] `https://www.data.gouv.fr/fr/datasets/elections-departementales-2015-resultats-par-bureaux-de-vote/`

[34] `https://www.bundeswahlleiter.de/en/bundestagswahlen/BTW_BUND_13/ergebnisse/wahlkreisergebnisse/index.html`

United Kingdom I obtained information on the vote share of the United Kingdom Independence Party (UKIP), British Democrats, British National Party, Liberty GB party, and the National Front party in the United Kingdom's 2015 General Elections from the country's Electoral Commission website.[35] For each constituency, I calculated the percent vote share for these parties.

Belgium I downloaded voting results from the 2014 Belgian Federal Elections at the municipality level from the country's Election Board website.[36] I calculated the vote share for Vlaams Belang for each constituency.

Socioeconomic data

France I obtained data on unemployment, share of foreigners, number of asylum seeker centers, and population size from the National Institute of Statistic and Economic Studies (INSEE).

1. *Unemployment (2011).* Unemployment at the municipality level the 2011 census.[37]
2. *Share of foreigners (2011).* The share of non-nationals in each municipality from the 2011 census.[38]
3. *Asylum seekers (2014).* The number of asylum seeker centers in each municipality as of 2014.[39]
4. *Population (2011).* Population size in each municipality from the 2011 census.[40]

[35] `http://www.electoralcommission.org.uk/our-work/our-research/electoral-data`

[36] `http://www.elections.fgov.be/index.php?id=3265&L=1`

[37] `http://www.insee.fr/fr/themes/detail.asp?reg_id=99&ref_id=td-population-13`

[38] `http://www.insee.fr/fr/themes/detail.asp?reg_id=99&ref_id=td-nationalite-13`

[39] `http://www.insee.fr/fr/themes/detail.asp?reg_id=99&ref_id=equip-serv-action-sociale`

[40] `http://www.insee.fr/fr/themes/detail.asp?reg_id=99&ref_id=td-population-13`

Germany I downloaded data on unemployment, immigration, asylum seeker benefit receivers, and population size at the municipality level from The Regional Database Germany.[41] In order to access the data, it is necessary to create an account. Thus, I provide the names of the tables that I downloaded from the database.

1. *Unemployment (2015).* Unemployed individuals by selected groups of persons (Arbeitslose nach ausgewählten Personengruppen)
2. *Share of foreigners (2014).* Immigration and emigration by gender and age groups, over municipal boundaries, yearly total (Zu- und Fortzüge nach Geschlecht und Altersgruppen, über Gemeindegrenzen, jahressumme)
3. *Asylum seeker benefits receivers (2014).* Recipients of asylum seekers standard benefits, by gender, type of service, and age groups (Empfänger von Asylbewerberregelleistungen, Geschlecht, Art der Leistung, Altersgruppen)
4. *Population size (2011).* Population size at the municipality level from the 2011 census.

United Kingdom I obtained data from the 2011 census on unemployment, immigration, population size, religion, and ethnicity at the level of the Mid-layer super output area (MSOA), which is roughly equal to the size of a neighborhood, from the United Kingdom's Office of National Statistics.[42] I provide the names and numbers of tables that I downloaded from the database.

1. *Unemployment (2011).* KS601UK – Economic activity
2. *Share of foreigners (2011).* QS803EW – length of residence in the UK
3. *Population (2011).* KS101EW – Usual resident population
4. *Religion (2011).* LC1202EW – Household composition by religion of Household Reference Person (HRP)
5. *Ethnic group (2011).* KS201EW – Ethnic group

[41] `https://www.regionalstatistik.de/genesis/online/online;jsessionid=EE45147898822814978BE734145275C4?operation=sprachwechsel&option=en`

[42] `https://www.ons.gov.uk`

Belgium I downloaded data on unemployment, immigration, and population at the statistical sector (sub-municipality) level from the 2011 Belgian census.[43] I provide the names of the tables that I downloaded from the database.

1. *Unemployment (2011).* Employed population by gender and age group - Total population - Statistical Sector (Werkende bevolking naar geslacht en leeftijdsklasse - Totale bevolking - Statistische sector)
2. *Share of foreigners, population (2011).* Population of Belgian and foreign nationality by gender – Statistical sector (Bevolking van Belgische en vreemde nationaliteit naar geslacht - Statistische sector)

Shape files

France I obtained shape files for the electoral cantons in France's 2015 Departmental Elections from the country's open platform of public data.[44]. For other administrative data, I obtained shape files of the contours of France's municipalities from France's open platform for public data.[45]

Germany I downloaded shape files of electoral constituencies in the 2013 German Federal Elections from Germany's Federal Returning Officer's Office.[46] For other socioeconomic variables, I used shape files from the contours of Germany's administrative boundaries.[47]

United Kingdom I obtained shape files for U.K. parliamentary constituencies from *MapIt*, a charity that provides data on contours of administrative areas in the United

[43] `http://census2011.fgov.be/download/statsect_nl.html`

[44] `https://www.data.gouv.fr/fr/datasets/contours-des-cantons-electoraux-departementaux-2015/`

[45] `https://www.data.gouv.fr/fr/datasets/geofla-communes/`

[46] `https://www.bundeswahlleiter.de/en/bundestagswahlen/BTW_BUND_13/wahlkreiseinteilung/kartographische_darstellung.html`

[47] `https://www.zensus2011.de/DE/Infothek/Begleitmaterial_Ergebnisse/Begleitmaterial_node.html`

Kingdom.[48] I then matched the constituency-level vote share of far-right parties to the relevant polygon. For census data at the MSOA level, I used shape files from the Office of National Statistics.[49]

Belgium I downloaded the shape files of the contours of Belgium's statistical sectors (sub-municipality level) from Statistics Belgium, the official website of national statistics.[50]

[48] `https://mapit.mysociety.org/areas/WMC.html`

[49] `http://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/guide-method/geography/products/census/spatial/2011/index.html`

[50] `http://statbel.fgov.be/nl/statistieken/opendata/datasets/tools/geografisch/`

2.6.6 Social media usage by ISIS activists in the United States

Table 2.25 provides details on the social media usage of over a hundred of individuals charged in the United States with providing material support for ISIS or plotting a violent attack on the organization's behalf. Data come from criminal complains filed against these individuals in United States courts, which describe in detail these individuals' pro-ISIS activities. I coded each case according to whether the individual used social media platforms such as Twitter or Facebook during their radicalization process. In addition, I documented whether the individual expressed publicly his or her support for the Islamic State and its ideology. Understanding whether radicalizing individual post *public* social media posts is important for this paper's data collection method, which assumes that it is possible to observe (at least part of) one's radicalization process by scraping information on his or her online behavior. The data show that the majority of these individuals used social media when radicalizing (about 62%). Among those who used social media, the vast majority (about 86%) posted publicly their support for ISIS.

Table 2.25: Social media usage by ISIS supporters in the United States

	Name	Location	Used social media	Posted public posts
1	Samy el-Goarany	New York	1	1
2	Ahmed Mohammed El Gammal	Arizona	1	1
3	Abdul Malik Abdul Kareem	Phoenix, AZ	0	0
4	Elton Francis Simpson	Phoenix, AZ	1	1
5	Nader Ehuzayel	Santa Ana, California	1	1
6	Muhanad Badawi	Santa Ana, California	1	1
7	Nicholas Michael Teausant	Acampo, CA	1	1
8	Adam Dandach	Orange County, CA	0	0
9	Enrique Marquez Jr.	Riverside, CA	0	0
10	Aws Mohammed Younis al-Jayab	Sacramento, CA	1	0
11	Mahamad Saeed Koadimati	San Diego, CA	1	0
12	Shannon Maureen Conley	Denver, CO	1	0
13	James Gonzalo Medina	Hollywood, FL	0	0
14	Harlem Suarez	Key West, FL	1	1
15	Gregory Hubbard	West Palm Beach, FL	1	0
16	Dayne Antani Christian	Lake Park, FL	0	0
17	Darren Arness Jackson	West Palm Beach, FL	0	0
18	Miguel Moran Diaz	Miami-Dade, FL	1	1
19	Robert B. Jackson	Pensacola, FL	1	1
20	Leon Nathan Davis	Augusta, GA	0	0
21	Hasan R. Edmonds	Aurora, IL	1	1
22	Jonas M. Edmonds	Aurora, IL	0	0
23	Mhammed Hamzah Khan	Bolingbrook, IL	0	0
24	Ramiz Zijad Hodzic	Saint Louis, MO	1	1
25	Sedina Unkic Hodzic	Saint Louis, MO	1	1
26	Nihad Rosic	Utica, NY	1	1
27	Mehida Medy Salkicevic	Schiller Park, IL	1	1
28	Armin Harcevic	Saint Louis, MO	1	1
29	Jasminka Ramic	Rockford, IL	1	1
30	Abdullah Ramo Pazara	Saint Louis, MO	1	0
31	Akrami I. Musleh	Brownsburg, IN	1	1
32	Alexander E. Blair	Topeka, KS	0	0
33	John T. Booker	Topeka, KS	1	1
34	Alexander Ciccolo	Adams, MA	1	1
35	David Wright	Everett, MA	0	0
36	Mohamed Elshinaway	Edgewood, MD	1	1
37	Khalil Abu Rayyan	Dearborn Heights, MI	1	1
38	Sebastian Gregerson	Detroit, MI	0	0
39	Al-Hamzah Mohammad Jawad	East Lansing, MI	0	0
40	Abdirizak Mohamed Warsame	Eagan, MN	0	0
41	Abdul Raheem Habil Ali-Skelton	Glencoe, MN	0	0
42	Mohamed Abdihamid Farah	Minneapolis, MN	0	0
43	Adnan Abdihamid Farah	Minneapolis, MN	1	1
44	Abdurahman Yasin Daud	Minneapolis, MN	0	0
45	Zacharia Yusuf Abdurahman	Minneapolis, MN	0	0
46	Hanad Mustafe Musse	Minneapolis, MN	0	0
47	Guled Ali Omar	Minneapolis, MN	0	0
48	Hamza Ahmed	Minneapolis, MN	1	1
49	"H.A.M"	Burnsville, MN	1	1
50	Abdullahi Yusuf	Inver Grove Heights, MN	1	1
51	Abdi Nur	Minneapolis, MN	1	1
52	Yusra Ismail	St. Paul, MN	0	0
53	Safya Roe Yassin	Bolivar, MO	1	1
54	Jaelyn Delshaun Young	Starkville, MS	1	1
55	Muhammad Oda Dakhlalla	Starkville, MS	0	0

Note: The table provides details on the social media usage d of individuals charged in the United States with providing material support for ISIS or plotting a violent attack on the organization's behalf. Data come from criminal complains filed against these individuals in United States courts, which describe in detail these individuals' pro-ISIS activities.

Social media usage by ISIS supporters in the United States

	Name	Location	Used social media	Posted public posts
56	Justin Nojan Sullivan	Burke County, NC	1	0
57	Erick Jamal Hendricks	Charlotte, NC	1	1
58	Avin Marsalis Brown	Raleigh, NC	0	0
59	Akba Johad Jordan	Raleigh, NC	0	0
60	Donald Ray Morgan	Rowan County, NC	1	1
61	Nader Saadeh	Rutherford, NJ	1	1
62	Alaa Saadeh	West New York, NJ	0	0
63	Samuel Rahamin Topaz	Fort Lee, NJ	1	1
64	Tairod Nathan Webster Pugh	Neptune, NJ	0	0
65	Sajmir Alimehmeti	Bronx, NY	1	0
66	Abdursasul Hasanovich Juraboev	Brooklyn, NY	1	1
67	Akhror Saidakhmetov	Brooklyn, NY	1	1
68	Arbor Habibov	Brooklyn, NY	0	0
69	Dilkhayot Kasimov	Brooklyn, NY	0	0
70	Almal Zakirov	Brooklyn, NY	0	0
71	Mohimanul Bhuiya	Brooklyn, NY	0	0
72	Noelle Velentzas	Queens, NY	0	0
73	Asia Siddiqui	Queens, NY	1	1
74	Arafat M. Nagi	Lackawanna, NY	1	1
75	Ali Saleh	Fort Wayne, IN	1	1
76	Munther Omar Saleh	Queens, NY	1	1
77	Emanuel L. Luchtman	Rochester, NY	1	0
78	Mufid A. Elfgeeh	Rochester, NY	1	1
79	Farred Mumuni	Staten Island, NY	0	0
80	Terrence Joseph Mcneil	Akron, OH	1	1
81	Christopher Lee Cornell	Cincinnati, OH	1	1
82	Amir Aid Abdul Rahman Al-Ghazi	Sheffield Lake, OH	1	1
83	Munir Abdulkader	West Chester, OH	1	1
84	Jalil Ibn Amer Aziz	Harrisburg, PA	1	1
85	Keonna Thomas	Philadelphia, PA	1	1
86	David Wright	Everett, MA	0	0
87	Nicholas Rovinski	Warwick, RI	1	1
88	Usama Rahim	Roslindale, MA	0	0
89	Michael Todd Wolfe	Houston, TX	0	0
90	Omar Faraj Saeed Al Hardan	Houston, TX	0	0
91	Asher Abid Khan	Spring, TX	1	0
92	Sixto Ramiro Garcia	Houston, TX	1	1
93	Bilal Abood	Mesquite, TZ	1	1
94	Mohamad Jamal Khweis	Alexandria, VA	1	1
95	Haris Qamar	Burke, VA	1	1
96	Nicholas Young	Fairfax, VA	0	0
97	Amine El Khalifi	Fairfax, VA	1	1
98	Yusuf Abdirizak Wehelie	Failfax, VA	0	0
99	Heather Elizabeth Coffman	Richmond, VA	1	1
100	Mohamed Bailor Jalloh	Sterling, VA	1	1
101	Ali Shukri Amin	Woodbridge, VA	1	1
102	Joseph Hassan Farrokh	Woodbridge, VA	0	0
103	Mhamoud Amin Mohamed Elhassan	Woodbridge, VA	0	0
104	Daniel Seth Franey	Montesano, WA	1	1
105	Joshua Van Haften	Madison, WI	1	1
Proportion using social media			0.62	
Proportion posting public posts (among those using social media)				0.86

Note: The table provides details on the social media usage d of individuals charged in the United States with providing material support for ISIS or plotting a violent attack on the organization's behalf. Data come from criminal complains filed against these individuals in United States courts, which describe in detail these individuals' pro-ISIS activities.

2.6.7 Additional results

Table 2.26: Western foreign fighters and online radicalization by country

	(1) Number of Twitter users flagged as ISIS activists	(2) Number of Twitter users posting highly radical content	(3) Number of ISIS accounts followed	(4) Number of Twitter users suspended from Twitter
Number of foreign fighters	0.132***	0.156	78.291***	0.280***
	(0.029)	(0.133)	(20.145)	(0.104)
Constant	11.702	3.422	4,733.557	42.460
	(15.558)	(71.227)	(10,748.790)	(55.287)
Population controls	✓	✓	✓	✓
Number of observations	46	46	46	46
R^2	0.392	0.337	0.433	0.350

Note: The table reports the correlation between online radicalization measures and foreign fighter counts in European countries, controlling for population size. It can be seen that all online radicalization variables are positively correlated with the number of foreign fighters in each country, with the number of users flagged as ISIS activists, number of ISIS accounts followed, and the number of users suspended from Twitter significant at the 5% level.
*p<0.1; **p<0.05; ***p<0.01

Table 2.27: Different cutoffs for classifying top posters of radical content

	(1) Top 5%	(2) Top 10%	(3) Top 15%	(4) Top 20%	(5) Top 25%
Far-right vote share (%)	0.99***	0.87*	0.99***	0.99***	2.04
	(0.33)	(0.50)	(0.33)	(0.33)	(1.26)
Unemployment (%)	1.15	3.02**	1.15	1.15	8.10**
	(0.76)	(1.26)	(0.76)	(0.76)	(3.54)
Foreigners (%)	0.42	0.42	0.42	0.42	-1.26
	(0.29)	(0.46)	(0.29)	(0.29)	(1.09)
Constant	43.67**	73.52***	43.67**	43.67**	283.37***
	(16.97)	(23.71)	(16.97)	(16.97)	(56.36)
Population controls	✓	✓	✓	✓	✓
Country fixed effects	✓	✓	✓	✓	✓
R^2	0.002	0.002	0.002	0.002	0.004
Number of clusters	2,655	2,655	2,655	2,655	2,655
Number of observations	112,254	112,254	112,254	112,254	112,254

Robust standard errors in parentheses, clustered at the locality level. Base country is Belgium. All coefficients are $\times$ 1,000 to account for the skewed distribution of the dependent variables.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.28: Correlates of flagged accounts

	Pr(flagged as ISIS account)					
	(1)	(2)	(3)	(4)	(5)	(6)
Sympathy with ISIS (# tweets)	0.002*** (0.00001)					
Travel to Syria or foreign fighters (# tweets)		0.001*** (0.00001)				
Life in ISIS territories (# tweets)			0.001*** (0.00001)			
Anti-West (# tweets)				0.002*** (0.00002)		
Syrian war (# tweets)					0.001*** (0.00002)	
Islam (# tweets)						0.001*** (0.00001)
Constant	0.008*** (0.0001)	0.008*** (0.0001)	0.010*** (0.0001)	0.008*** (0.0001)	0.011*** (0.0001)	0.006*** (0.0001)
Observations	1,052,842	1,052,842	1,052,842	1,052,842	1,052,842	1,052,842
R^2	0.018	0.020	0.008	0.016	0.003	0.038

*p<0.1; **p<0.05; ***p<0.01

Table 2.29: Socioeconomic correlates of support for ISIS on Twitter

	(1) Indicator for top 1% radical content	(2) Indicator for top 1% radical content, without Islam	(3) Flagged as an ISIS activist	(4) Suspended from Twitter	(5) Number of ISIS accounts following
Far-right vote share (%)	0.25***	0.25***	0.30***	0.09	86.48***
	(0.06)	(0.06)	(0.04)	(0.13)	(15.04)
Unemployment (%)	0.20	0.18	-0.20*	-1.24***	-111.71***
	(0.18)	(0.18)	(0.12)	(0.32)	(30.33)
Foreigners (%)	0.10*	0.11*	0.26***	-0.06	84.01***
	(0.06)	(0.06)	(0.04)	(0.12)	(15.57)
Constant	8.57**	7.68**	-9.76***	35.07***	1116.19
	(3.71)	(3.67)	(1.91)	(6.69)	(739.80)
Population controls	✓	✓	✓	✓	✓
Country fixed effects	✓	✓	✓	✓	✓
R^2	0.0003	0.0003	0.006	0.002	0.006
Number of observations	112,254	112,254	112,254	112,250	112,250

Heteroskedasticity robust standard errors in parentheses. Base country is Belgium.
All coefficients are × 1,000 to account for the skewed distribution of the dependent variables.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.30: Socioeconomic correlates of posting pro-ISIS content on Twitter

	(1) Sympathy with ISIS	(2) Life in ISIS territories	(3) Travel to Syria or foreign fighters	(4) Syrian war	(5) Anti-West	(6) Islam
Far-right vote share (%)	0.05***	0.07***	0.07***	0.04***	0.04***	0.08***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
Unemployment (%)	0.12***	0.14***	0.16***	0.11***	0.13***	0.23***
	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.04)
Foreigners (%)	0.02***	0.02**	0.03***	0.02**	0.02**	0.04**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
Constant	3.51***	5.53***	5.66***	2.68***	3.22***	7.49***
	(0.43)	(0.57)	(0.61)	(0.35)	(0.42)	(0.86)
Population controls	✓	✓	✓	✓	✓	✓
Country fixed effects	✓	✓	✓	✓	✓	✓
R^2	0.002	0.005	0.003	0.003	0.003	0.001
Number of observations	112,254	112,254	112,254	112,254	112,254	112,254

Heteroskedasticity robust standard errors in parentheses. Base country is Belgium.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.31: Unemployed immigrants and support for ISIS on Twitter

	(1) Indicator for top 1% radical content	(2) Indicator for top 1% radical content, without Islam	(3) Number of ISIS accounts following	(4) Flagged as an ISIS activist	(5) Suspended from Twitter
Unemployed immigrants (%)	0.09	0.11	0.30	-0.47	177.55**
	(0.36)	(0.35)	(0.21)	(0.69)	(76.78)
Constant	14.89***	14.38***	0.43	26.64***	3990.51***
	(2.71)	(2.67)	(0.78)	(3.67)	(257.66)
Population controls	✓	✓	✓	✓	✓
Country fixed effects	✓	✓	✓	✓	✓
R^2	0.0001	0.0001	0.003	0.001	0.001
Number of observations	90,516	90,516	90,516	90,514	90,514

Heteroskedasticity robust standard errors in parentheses. Base country is Belgium.
All coefficients are × 1,000 to account for the skewed distribution of the dependent variables.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.32: Asylum seekers and support for ISIS on Twitter

	(1) Indicator for top 1% radical content	(2) Indicator for top 1% radical content, without Islam	(3) Number of ISIS accounts following	(4) Flagged as an ISIS activist	(5) Suspended from Twitter
Asylum seekers (%, sd units)	0.12	0.07	-5.04***	-4.03***	-675.54***
	(0.68)	(0.68)	(0.53)	(1.31)	(182.02)
Constant	5.96***	5.34***	14.12***	50.98***	6852.84***
	(2.05)	(2.00)	(2.05)	(4.63)	(794.64)
Controls	✓	✓	✓	✓	✓
Country fixed effects	✓	✓	✓	✓	✓
R^2	0.0001	0.0001	0.005	0.001	0.001
Number of observations	88,388	88,388	88,388	88,386	88,386

Heteroskedasticity robust standard errors in parentheses. Base country is Germany.
All coefficients are × 1,000 to account for the skewed distribution of the dependent variables.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.33: Socioeconomic correlates of ISIS Twitter activism in Western countries (By country)

	(1) Indicator for top 1% radical content	(2) Indicator for top 1% radical content, without Islam	(3) Flagged as an ISIS activist	(4) Suspended from Twitter	(5) Number of ISIS accounts following
		France			
Far-right vote share (%)	0.25**	0.25**	1.01***	1.45***	544.75***
	(0.10)	(0.10)	(0.11)	(0.23)	(39.30)
Unemployment (%)	0.16	0.06	-2.55***	-6.10***	-2161.42***
	(0.45)	(0.44)	(0.45)	(1.08)	(199.38)
Foreigners (%)	0.01	0.00	0.38***	1.35***	526.10***
	(0.16)	(0.16)	(0.14)	(0.35)	(55.70)
Constant	0.77	0.69	-16.08***	29.63***	-1801.84**
	(2.84)	(2.82)	(2.75)	(6.43)	(836.50)
Population controls	✓	✓	✓	✓	✓
R^2	0.001	0.001	0.020	0.005	0.034
Number of observations	24,829	24,829	24,829	24,829	24,829
		Germany			
Far-right vote share (%)	0.64	0.36	16.14***	14.31***	7274.78***
	(1.61)	(1.64)	(3.58)	(4.53)	(1869.22)
Unemployment (%)	-2.03	-1.05	-16.21***	-12.34**	-236.56
	(2.35)	(1.40)	(2.50)	(5.65)	(1268.02)
Foreigners (%)	0.25	0.38*	2.84***	4.71***	887.04
	(0.30)	(0.20)	(0.48)	(1.08)	(570.09)
Constant	9.17	0.53	-51.60**	-51.61	-44764.9**
	(19.72)	(12.72)	(24.79)	(49.98)	(17643.41)
Population controls	✓	✓	✓	✓	✓
R^2	0.001	0.002	0.028	0.016	0.010
Number of observations	5,544	5,544	5,544	5,543	5,543

Heteroskedasticity robust standard errors in parentheses.
All coefficients are × 1,000 to account for the skewed distribution of the dependent variables.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Socioeconomic correlates of ISIS Twitter activism in Western countries (By country) - Cont.

	(1) Indicator for top 1% radical content	(2) Indicator for top 1% radical content, without Islam	(3) Flagged as an ISIS activist	(4) Suspended from Twitter	(5) Number of ISIS accounts following
		Belgium			
Far-right vote share (%)	-1.68	-1.44	-0.11	4.76**	671.31***
	(1.82)	(1.80)	(0.10)	(2.36)	(152.08)
Unemployment (%)	-1.22	-1.46	-1.66	0.84	830.00***
	(2.15)	(2.13)	(1.18)	(4.88)	(192.90)
Foreigners (%)	0.21	0.18	0.07	-0.01	58.43***
	(0.30)	(0.30)	(0.05)	(0.34)	(12.89)
Constant	8.99	10.28	6.39	60.72	-5458.01***
	(12.94)	(12.88)	(4.59)	(45.84)	(1390.37)
Population controls	✓	✓	✓	✓	✓
R^2	0.002	0.001	0.005	0.006	0.037
Number of observations	1,821	1,821	1,821	1,821	1,821
		United Kingdom			
Far-right vote share (%)	0.35***	0.39***	-0.05	-1.17***	-142.88***
	(0.10)	(0.11)	(0.05)	(0.21)	(20.21)
Unemployment (%)	-0.09	-0.09	-0.33***	-0.70*	-208.64***
	(0.28)	(0.28)	(0.08)	(0.41)	(25.35)
Foreigners (%)	0.12	0.17	-0.04	-0.92***	-102.45***
	(0.10)	(0.10)	(0.06)	(0.20)	(20.98)
Constant	-11.22**	-12.32**	-0.87	105.48***	4419.37***
	(5.16)	(5.16)	(2.24)	(9.37)	(804.03)
Population controls	✓	✓	✓	✓	✓
R^2	0.0004	0.0004	0.0002	0.001	0.001
Number of observations	80,060	80,060	80,060	80,057	80,057

Heteroskedasticity robust standard errors in parentheses.
All coefficients are × 1,000 to account for the skewed distribution of the dependent variables.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.34: Salient and non-salient radicalization: Online ISIS support and Terrorist attacks (France)

Dependent variable: Number of terrorist attacks in a locality	(1)	(2)	(3)	(4)
Number of users posting highly radical content	−0.001***			
	(0.0002)			
Number of users flagged as ISIS activists		−0.002***		
		(0.0002)		
Number of users suspended from Twitter			−0.001***	
			(0.0001)	
Number of ISIS accounts followed				−0.00000***
				(0.00000)
Constant	−0.0003**	−0.0003***	−0.0003***	−0.0003***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Population controls	✓	✓	✓	✓
Number of observations	32,676	32,676	32,676	32,676
R^2	0.475	0.478	0.477	0.477

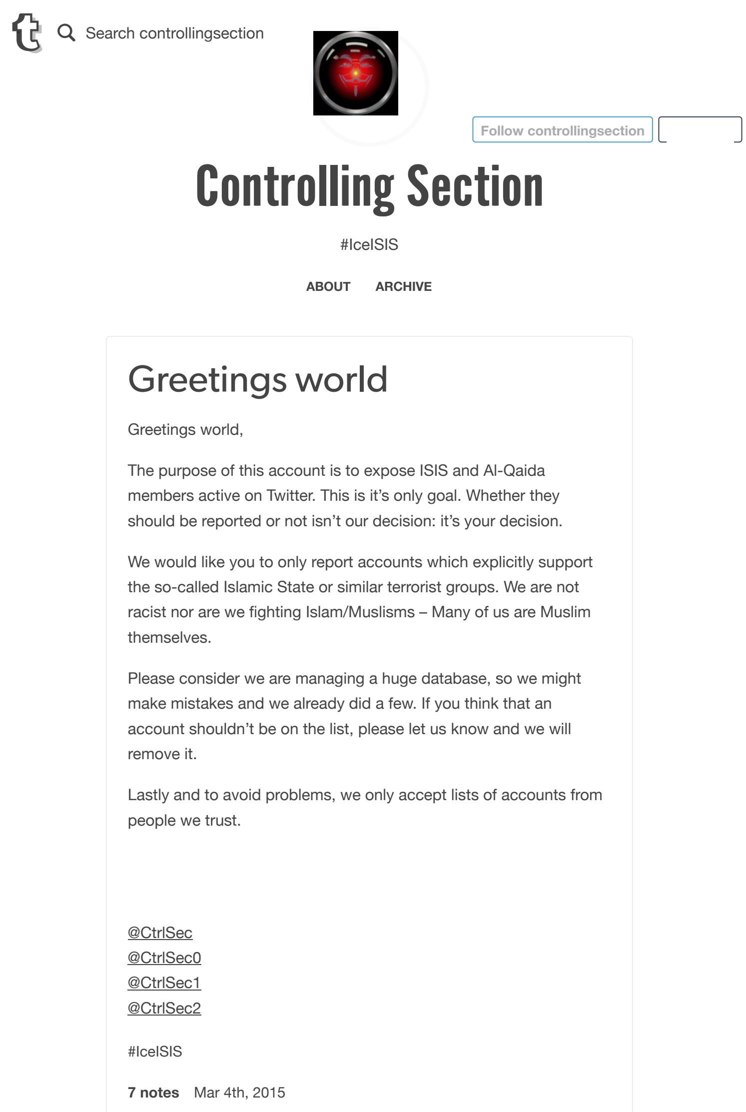
Note: *p<0.1; **p<0.05; ***p<0.01.
Unit of analysis is the municipality.

Table 2.35: Salient and non-salient radicalization: Online ISIS support and Terrorist attacks (UK)

Dependent variable: Number of terrorist attacks in a locality	(1)	(2)	(3)	(4)
Number of users posting highly radical content	−0.00000			
	(0.00003)			
Number of users flagged as ISIS activists		−0.00000		
		(0.0001)		
Number of users suspended from Twitter			−0.00000	
			(0.00001)	
Number of ISIS accounts followed				−0.000
				(0.00000)
Constant	0.0003	0.0003	0.0003	0.0003
	(0.003)	(0.003)	(0.003)	(0.003)
Population controls	✓	✓	✓	✓
Number of observations	7,201	7,201	7,201	7,201
R^2	0.0001	0.0001	0.0001	0.0001

Note: *p<0.1; **p<0.05; ***p<0.01
Unit of analysis is the MSOA.

Figure 2.19: @CtrlSec request to expose ISIS members on Twitter



Search controllingsection

Follow controllingsection

Controlling Section

#IceISIS

ABOUT ARCHIVE

Greetings world

Greetings world,

The purpose of this account is to expose ISIS and Al-Qaida members active on Twitter. This is it's only goal. Whether they should be reported or not isn't our decision: it's your decision.

We would like you to only report accounts which explicitly support the so-called Islamic State or similar terrorist groups. We are not racist nor are we fighting Islam/Muslisms – Many of us are Muslim themselves.

Please consider we are managing a huge database, so we might make mistakes and we already did a few. If you think that an account shouldn't be on the list, please let us know and we will remove it.

Lastly and to avoid problems, we only accept lists of accounts from people we trust.

@CtrlSec
@CtrlSec0
@CtrlSec1
@CtrlSec2

#IceISIS

7 notes Mar 4th, 2015

Source: `http://controllingsection.tumblr.com/post/112703617620/greetings-world`

Figure 2.20: Example of @CtrlSec real-time flagging of ISIS acounts

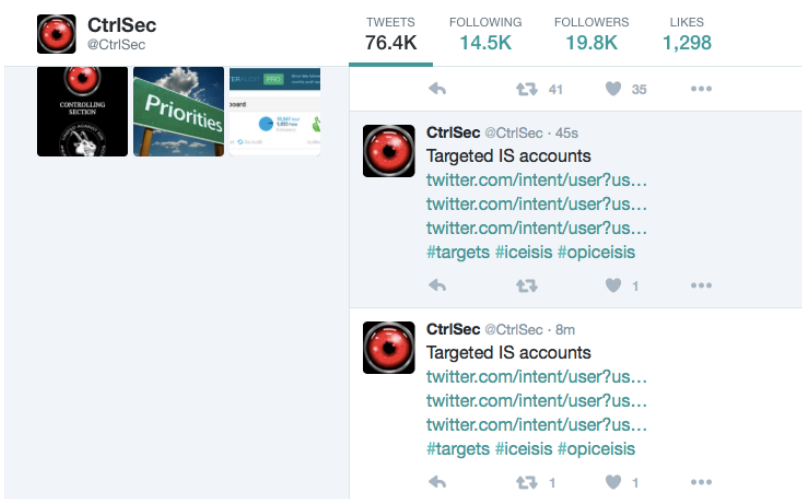
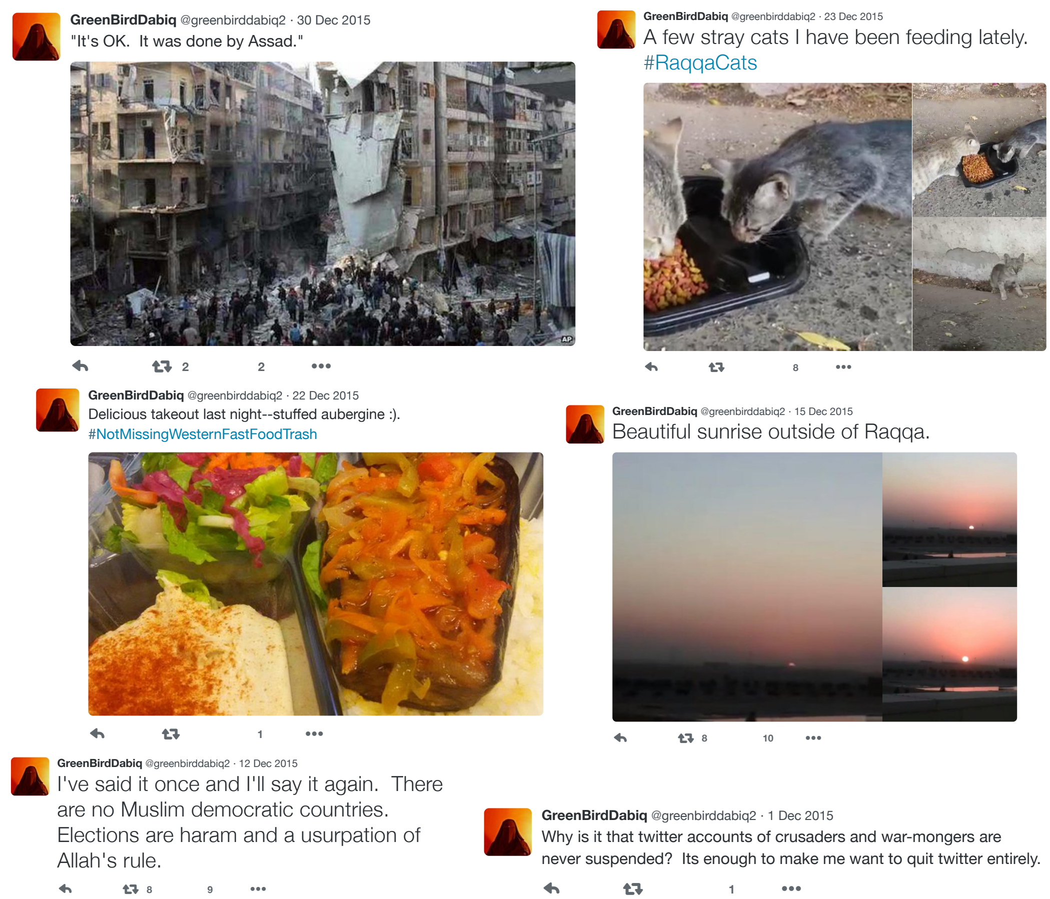


Figure 2.21: Example of ISIS accounts

Figure 2.22: Example of a Western fighter tweeting from Syria



Note: This account has already been suspended as of February 2016.

Figure 2.23: Example of a suspended account

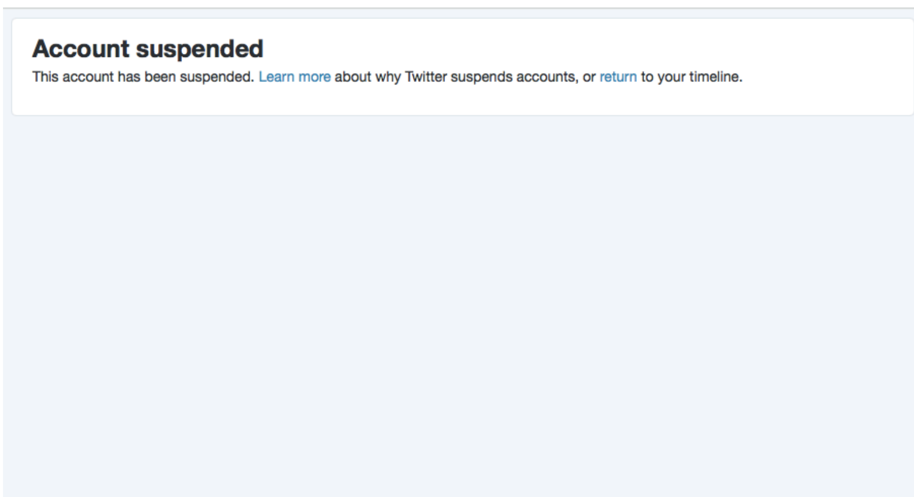
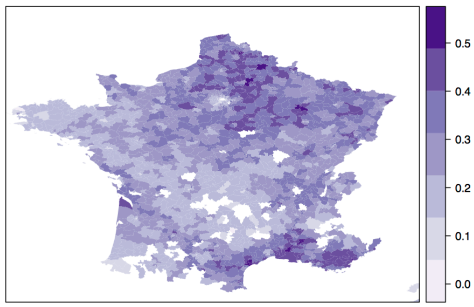
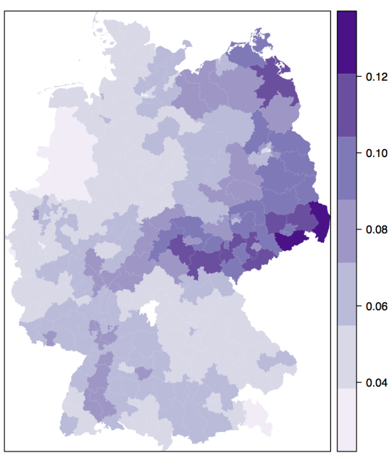


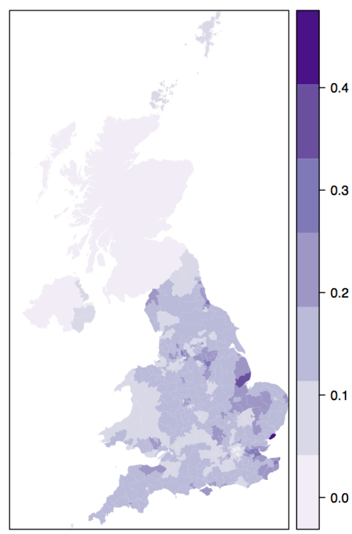
Figure 2.24: Vote share for far-right parties



(a) France



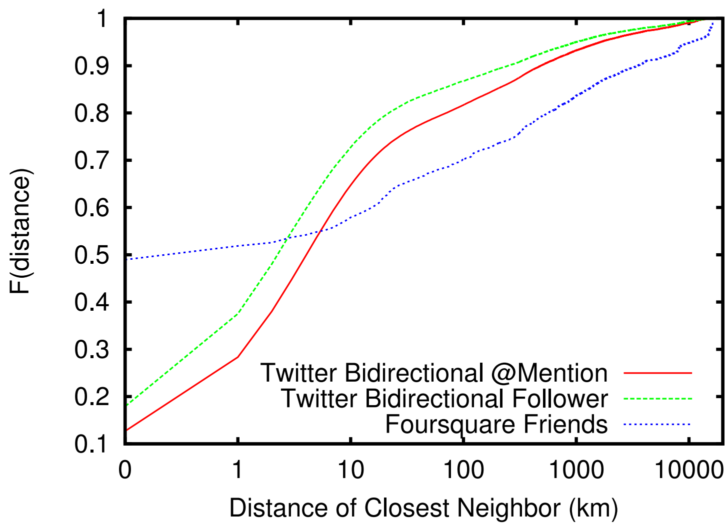
(b) Germany



(c) United Kingdom

Note: For France, the map displays the vote share for the Front National party in the 2015 departmental elections at the electoral canton level. For Germany, the map displays the vote share for the Alternative for Germany (AfD) party and the National Democratic Party (NPD) in the 2013 federal elections. For the UK, the map represents the vote share for the British Democrats, British National Party, Liberty GB party, National Front party, and United Kingdom Independence Party in the 2015 U.K. parliamentary general elections. See Table 2.4 for details on data sources.

Figure 2.25: The cumulative distribution functions for the distance to a user's geographically closest friend (Figure taken from Jurgens (2013))



Note: The figure, taken from the study of Jurgens (2013), shows cumulative distribution functions (CDFs) of users' geographical distance to their closest neighbor in three social media networks. In the figure, the x axis shows disance in kilometers, and the y axis shows the probability that the closest neighbor for each user is located x distance or less from that user. It can be seen that more than half of the users in these three networks had neighbors that were located within 4 kilometers from them, thereby allowing location prediction within 4-kilometer bounds.

Figure 2.26: Tweet content classification task instructions for CrowdFlower workers

Classify Syrian Civil War Tweets (English)

Instructions ▴

Please label each tweet by checking all labels that correctly describe its content. If a tweet does not fit any of the labels, check "None of the Above".

Category	Description
Anti-West	Anti-West rhetoric, criticizing Western countries' foreign policy and military operations in the Middle East
Islamic faith	Expressions of faith in the Islamic religion, Islamic quotes, and prayers and/or requests for prayers
IS sympathy	Expressions of support or sympathy with the Islamic State, its ideology and its activities in territories under its control
Life in IS territories	Tweets from Islamic State activists describing their life in the territories controlled by the Islamic State; includes descriptions of daily activities under Islamic State rule, fighting; things that 'market' the life in Syria to potential foreign fighters
Travel to Syria / foreign fighters	Tweets describing interest or intent to travel to Syria, and/or discussion of foreign fighters
Syrian war	Tweets describing events in the Syrian civil war and/or discussion/analysis of those events
Islamophobia	Tweets describing unfair treatment of Muslims and/or discrimination against Muslims in non-Muslim majority countries

Islam is not a religion as Christianity/Judaism nor a political belief as Capitalism/Communism but rather it is a comple…

Classification:

- ☐ Anti-West
- ☐ Islamic faith
- ☐ IS sympathy
- ☐ Life in IS territories
- ☐ Travel to Syria / foreign fighters
- ☐ Syrian war
- ☐ Islamophobia
- ☐ None of the Above

UK extremist's sharia law photo used in free speech ad

Classification:

- ☐ Anti-West
- ☐ Islamic faith
- ☐ IS sympathy
- ☐ Life in IS territories
- ☐ Travel to Syria / foreign fighters
- ☐ Syrian war
- ☐ Islamophobia
- ☐ None of the Above

Note: This is an example of a CrowdFlower task to classify English language tweets on various dimensions. Classified tweets are included in a training set to predict the content of unclassified tweets. The classification was carried out in English, French, Arabic, and German.

Table 2.36: Top 50 words across topics

	Topic	Top 50 words
1	Sympathy with ISIS	amp.will, beauti.islam, anyth.deserv, protector.htt, gambl, allah.muhammad, visceg, ideolog, bros, murthad, bomb.defeat, anonym.claim, day.let, allah.martyr, idni, attaref, bomb.raid, anyth.el, learn.quran, muwahideen, child.dont, antiislam, isi, frien, insignif, abysinnia, beer, back.start, otherhezbollat, aisha, belong.pious, incap, yet.dont, alway.week, ghazl, ago.near, hazima, week.follow, alfurqan, amp.realli, usrussia, bizarr, upcom, backbit.hurt, palestin.sha, clemenc, concern.islam, boy.gun, shepherd, attritionlos
2	Life in ISIS territories	samir, ate, besieg, nobl, dress, today.https, bomb.etc, iraqi.children, dua.syria, rt.propaganda, saa, cctv, masharialashwaq, jaish.alislam, machin, momineen, alloush.martyr, pkk.terrorist, islam.court, baqir, trade, charli, allegi.htt, behind.five, border.guard, assad.barrel, univ, civilian.say, outpost, eatabl, flay, dael, citi.fallen, jahiliyah, almohammad, conniv, aid.kuffar, center, hussari, qaeda, kiss, antiaq, pkk, reloc, sayyidi, judici, khaleefah, allow.sleep, like.year, ad.encrypt
3	Travel to Syria or foreign fighters	engag, dua.will, come.europ, join, elev, countrysid, martyrdom.op, sane, cemeteri, ahzamiyah, kasiki, fu, blogger, fighter, abar, aynisa, kashmiri, breakdown, foreign, shut, australian, syria, militia, kurdishwho, iraqi, martyrdom, batch.recruit, europ, braclet, age, australian.teen, bangladeshi, bangladeshi.blogger, muay, excut, fled, milit, fool.know, amp.islam, arraqqa, armi.right, get.today, graciou, pastpres, sustain, join.isi, amer, loot, egypt, najaf
4	Syrian war	syria, syrian, airstrik, regim, rebel, suspicion, traumat, russian, https, rt, will.face, hamid, erad, offici.tell, go.arrest, ghannam, militari, attack, againstassad, bewar, strike, aldagestani, dhawahiri, hom, onesyourselv, kiss, jet, jay, southern, seiz, typo, regiment, fight, latest.twitter, amp.bomb, commit, besieg.town, civilian, explo, chao.eastern, allow.enemi, mani.report, prior, build.center, heartbreak, bombard, missil, barrel, armi, enemi.get
5	Anti-West	afp.un, advic.us, bush, cant.invad, america, design, come.condemn, china.well, georg, gunmen, rpli, gwot, anyon.israel, erad, ampstop, militiasjihadist, real.terrorist, dua.ikhwan, punch, amp.tomorrow, org, aftermath.us, deplet, islamampth, usa, western, attempt.stab, punish, criminalis, belief.crusad, catastroph, clark, rehman, antiwar, repercuss, alli.usa, heyena, usback, pentagon, holocaust, obama, brother.say, brother.allah, american, washington, alli.nato, democraci, gay.gambia, bashar, penni
6	Islam	allah, quran, muslim, prophet, allaah, nabi, prayer, surah, muhammad, qayyim, rather, islam, idhnillah, sin, tagh, allh, hiatus, breast, shayt, fast, qurn, habit, dawah, spous, hijab, alrahman, religion, qualifi, alnubala, nur, ayede, hellfir, qur, therebi, all, almubarak, believ, haqq, btw, bless, inshaal, aha, islaam, foremost, faith, assahab, torah, brother.allahuakbar, execut.saudia, paradis

Figure 2.27: Anti-Muslim marches organized by PEGIDA across Europe

Note: Photos credit: *Radio Free Europe Radio Liberty* (2016) and Malm (2015)

Chapter 3

Terrorism as a Provocation Strategy: Transnational Terrorism and Anti-Muslim Hostility in the West

Tamar Mitts[1]

[1] I would like to thank Christopher Blattman, Jasper Cooper, Lindsay Dolan, Page Fortna, Grant Gordon, Macartan Humphreys, Sarah Khan, Summer Lindsey, Joshua Mitts, Suresh Naidu, Kunaal Sharma, Camille Strauss-Kahn, and Lauren Young for their advice and feedback on various stages of this project. This research was approved by the Columbia University Institutional Review Board under protocol IRB-AAAQ7657.

Abstract

Research on the strategies of terrorism argues that armed groups use terrorism to mobilize recruitment by provoking governments to overreact with violence against potential supporters. By focusing on government responses, existing work has largely overlooked the way in which terrorism can affect the behavior of targeted populations. This study tests the argument that acts of terror provoke targeted populations to become indiscriminately hostile to the crowd from which potential supporters can be drawn. Combining data on over thirty terrorist attacks by radical jihadists in Europe and the United States from 2010 to 2016 with high-frequency, panel data from Twitter, I show that individuals significantly increase anti-Muslim rhetoric after they are exposed to terrorist violence. This effect spikes immediately after attacks, decays over time, but remains significantly higher than pre-attack levels up to a month after the events. The findings also show that acts of terror resulting in more casualties have a stronger effect on anti-Muslim sentiment than attacks causing low number of victims.

3.1 Introduction

Terrorist attacks are used by violent organizations to obtain strategic goals. While many of these objectives, such as extracting government concessions, are often not reached (Fortna, 2015), acts of terrorism can be particularly successful in achieving intermediate goals like increasing mobilization and recruitment. A large body of research on the strategies of terrorism has shown that armed groups carry out terrorist attacks to provoke targeted governments to retaliate in a manner that facilitates support for the group. Expecting a strong state response and associated collateral damage, armed groups seek to exploit grievances caused by governments' retaliatory actions to mobilize popular support (Bueno de Mesquita and Dickson, 2007; De Figueiredo and Weingast, 2000; Rosendorff and Sandler, 2010).

Most research analyzing the provocation strategy of terrorism has focused on provoking the targeted government. By focusing on state responses to terrorism, research has assumed, rather than explained, the behavior of the targeted population. The prevalent presumption in many provocation models is that the targeted population will push the government to take strong actions against the group (Kydd and Walter, 2006; Lake, 2002). However, people experiencing terrorism may change their behavior in ways that can directly facilitate support for armed groups, regardless of government actions. As shown in the first chapter of the dissertation, popular hostility against Muslims can increase radicalization and support for the Islamic State among potential Western supporters.

In this chapter, I argue that armed groups strategically use terrorist attacks to manipulate the behavior of the targeted population. I test an important observable implication of this argument — i.e., that terrorist acts will negatively change the sentiment of the targeted population towards the community from which the group seeks to recruit supporters. Specifically, I focus on jihadi violence in Europe and the United States and examine whether acts of terrorism systematically increase popular hostility against Muslims. Literature on the sources of anti-Muslim attitudes in the West has pointed to economic competition and cultural differences as the drivers of hostility (Adida, Laitin and Valfort, 2016; Dancygier and Laitin, 2014). However, there is growing evidence that perceiving Muslims as security threats plays an important role in increasing anti-Muslim sentiment, especially in Western countries (Wike and Grim, 2010; Das et al., 2009).

This study examines how over thirty terrorist attacks perpetrated by radical jihadists in the West from 2010 to 2016 shape anti-Muslim attitudes among individuals in targeted countries. Unlike prior research that has focused on aggregate national or sub-national patterns of responses to terrorism in the West (Gould and Klor, 2014; Hanes and Machin, 2014; Kaushal, Kaestner and Reimers, 2007), this study examines how terrorist attacks shape the behavior of individual citizens. It therefore sheds light on the microfoundations of anti-Muslim hostility by looking into how people change their own rhetoric after exposure to terrorism. Combining information on the timing and location of terrorist attacks with micro-level, high frequency panel data from Twitter, I examine how online posts of people in targeted countries change in the days and weeks succeeding acts of terrorism.

For each attack, I collected data on random samples of thousands of individuals in targeted countries, obtaining information on what they posted on Twitter in the two months surrounding the attack. Using text-as-data tools, I created individual-level measures of anti-Muslim sentiment, which are based on the similarity of Twitter posts to an anti-Muslim vocabulary generated from all the tweets that include anti-Muslim hashtags such as #banmuslims, #IslamIsTheProblem', and #hatemuslims',[2] as well as all tweets generated by politicians from far-right parties who oppose Muslims in Europe.[3]

[2]The full list of hashtags is: #banislam, #banmuslims, #BanShariaLaw, #IslamIsTheProblem, #hatemuslims, #hateislam, #IslamIsEvil, #NoMuslimRefugees, #eradicateislam, #islamstupidity, #ExterminateIslam, #StopIslam, #BanSharia, #StopRapeJihad, #BanTheBurka, #noislam, #NoSharia, #islamoutnow.

[3]These politicians come from the following parties: Front National in France, NPD and AfD in Germany, British Democrats, British National Party, Liberty GB, National Front and UKIP in the UK, and Vlaams

Results from thirty-six attacks in twelve countries show that individuals systematically and significantly increase posting of anti-Muslim content on social media after exposure to terrorism. The effect spikes immediately after attacks, decays over time, but remains significantly higher than pre-attack levels up to a month after the events. The finding also reveal that the impact of terrorist attacks on anti-Muslim rhetoric is similar for individuals who already express hostility to Muslims before the attacks and those who do not. In other words, terrorism seems to lead individuals in targeted countries to express greater anti-Muslim hostility, regardless of whether they were already hostile before the attacks. Finally, results show that the impact of terrorist attacks on anti-Muslim hostility increases with the lethality of the attack: acts of terror generating more casualties have a stronger effect on anti-Muslim sentiment than attacks generating low number of victims.

The chapter contributes to existing research on the strategies of terrorism by empirically demonstrating that terrorist attacks might be strategically used not only to elicit government response, but also to provoke citizen behavior. The next section discusses existing theoretical work on the provocation strategy of terrorism, and drawing on research on inter-group relations and the legacies of terrorism in the West, explains how the provocation logic might be applied to the behavior of the targeted population.

3.2 Terrorism, provocation, and the targeted population

3.2.1 The provocation strategy of terrorism

Terrorist attacks — acts of violence against civilians carried out for political reasons[4] — are strategically used by armed groups to obtain various goals. Among the most common objectives are publicizing the group's cause, attriting the adversary, and provoking the

Belang in Belgium.

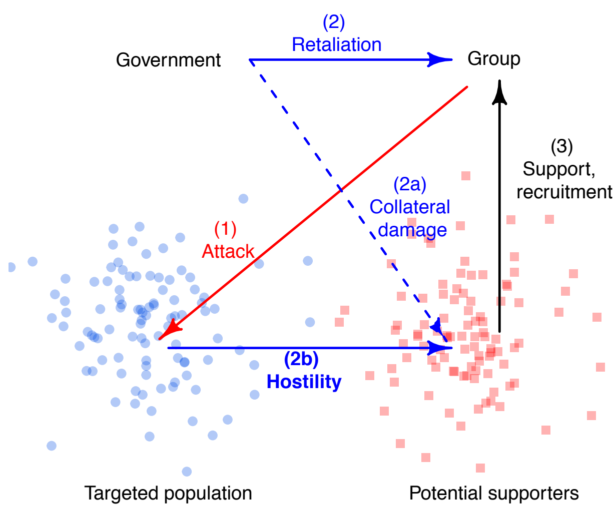
[4]Definition drawn from Merari (1993).

opponent to overreact with violence (Crenshaw, 1981; Kydd and Walter, 2006). Research studying the provocation strategy of terrorism, focusing primarily on the response of state actors, has argued that provocation is particularly useful for mobilizing supporters. According to theory, which is visually described in Figure 3.1, the logic works as follows. The armed group, seeking to attract supporters, attacks its adversary's civilians (Arrow 1). In response to the attack, the government retaliates with violence against the group (Arrow 2), which results in collateral damage affecting the aggrieved population (Arrow 2a). The damage inflicted by the government's retaliatory actions increases support from potential recruits (Arrow 3), as grievances from the state's retaliation frequently fit the group's recruitment narrative that the government is extreme and violent towards innocent civilians (Lake, 2002).

I argue that by focusing only on the behavior of targeted state actors, existing literature has ignored how terrorist attacks can be used to provoke the *targeted population* to respond in a manner conducive to rebel recruitment. Populations targeted by terrorism can play a central role in facilitating radicalization and support for armed groups, as citizens' hostile behavior towards potential supporters in day-to-day interactions amplifies and expands grievances caused by government responses to terrorism. In Chapter 2 of the dissertation, I showed that there is a strong relationship between measures of anti-Muslim hostility in Europe and individuals' likelihood of embracing the extremist ideology advanced by the Islamic State on social media. Armed organizations seeking to attract recruits might thus be able to manipulate levels of popular hostility against them by perpetrating acts of terrorism. This dynamic can be seen in Arrow 2b at the bottom of Figure 3.1.

In an essay published in the Islamic State's English-language magazine, *Dabiq*, the group stated that provoking hostility against Muslims with terrorist attacks in the West is among its top organizational goals. The group predicted that as a result of its violent acts, "Muslims in the West will quickly find themselves between one of two choices, they either

Figure 3.1: The provocation strategy of terrorism



apostatize to live amongst the kuffar [the unbelievers] ... or they perform hijrah [emigrate] to the Islamic State and thereby escape persecution from the crusader governments and citizens."[5] ISIS's recruitment strategy advocates a worldview in which there is an inherent division between Muslims and the West. By perpetrating acts of terrorism, the group hopes to perpetuate such perception among potential recruits using evidence of increased hostility against Muslims in the aftermath of terrorism.

3.2.2 Threat perception and anti-Muslim hostility in the West

Much of the literature on the strategies of terrorism has focused on group and state actions. Thus, it has not considered how terrorism might affect individual-level behavior among members of the targeted population. In order to understand how terrorist attacks can be strategically used to increase popular hostility against potential supporters, it is useful to consider insights from existing work on prejudice and intolerance, which examines how fear

[5]Source: https://ansarukhilafah.wordpress.com/2015/02/14/the-extinction-of-thr-grayzone/

and threat perception negatively impact inter-group attitudes and behavior.

A large body of work on the sources of inter-group hostility suggests that threat perception is one of the strongest predictors of exclusionary attitudes (Quillian, 1995; Stephan and Stephan, 1996; Sullivan, Piereson and Marcus, 1993). Focusing on democratic contexts, this research shows that people usually stay committed to liberal values and support tolerance and inclusion when in non-threatening conditions. However, when they feel threatened or experience inter-group conflict, people tend to compromise their commitments to democratic values, negatively stereotype out-group members, and increase hostile behavior (Marcus, 1995; Tajfel and Turner, 1979). This pattern has been found, for example, in the case of Israel, where Jewish Israeli citizens' attitudes toward Palestinian citizens became significantly more exclusionary when they felt greater levels of threat (Canetti-Nisim, Ariely and Halperin, 2008).

In the context of anti-Muslim animosity in the West, research has shown that much of the hostility against Muslims relates to perceived cultural differences and a sense of threat — rational or irrational — felt by the non-Muslim population. For example, Adida, Laitin and Valfort (2016) studied the failure of Muslims to integrate in France and found that taste-based discrimination, in which non-Muslims falsely equate Muslims with 'jihadists,' is an important driver of hostility. In survey work encompassing several countries in Europe, Wike and Grim (2010) found that perceiving Muslims as a security threat was the strongest predictor of anti-Muslim attitudes.

Given this pattern, terrorist attacks, which are known to increase fear and a sense of threat (Huddy et al., 2005; Becker, Rubinstein et al., 2004), can be a useful tool in the hands of groups seeking to manipulate levels of hostility against potential supporters. A large number of studies taking place after 9/11 found that individuals' threat perception significantly intensified prejudice and exclusionary attitudes (Canetti-Nisim, Ariely and Halperin, 2008; Coryn, Beale and Myers, 2004; Das et al., 2009; Fekete, 2004). Specifically, prejudice and hostile attitudes in Europe and the United States increased after 9/11 most strongly against Muslims and other minority groups with Middle Eastern heritage (Davila and Mora, 2005; Fetzer and Soper, 2005).

The goal of this study is to examine whether terrorist attacks perpetrated by armed groups adhering to jihadi ideology systematically, and across many contexts and events, increase hostility against Muslims in the West. If so, it provides support for the idea that groups might be using the provocation strategy not only against state governments, but also against targeted populations. Existing research has not been able to answer this question, because it either did not link individual behavior to actual violent events (Velasco González et al., 2008; Wike and Grim, 2010), or it focused on aggregate national or sub-national patterns almost exclusively around the 9/11 attacks in the United States (Åslund and Rooth, 2005; Gould and Klor, 2014).[6] This chapter seeks to contribute to existing work on the provocation strategy and the legacies of terrorism by analyzing a much larger number of jihadi-inspired attacks in the West, along with individual-level data on the behavior of targeted populations. The next section describes the data collection method and empirical strategy.

3.3 Data

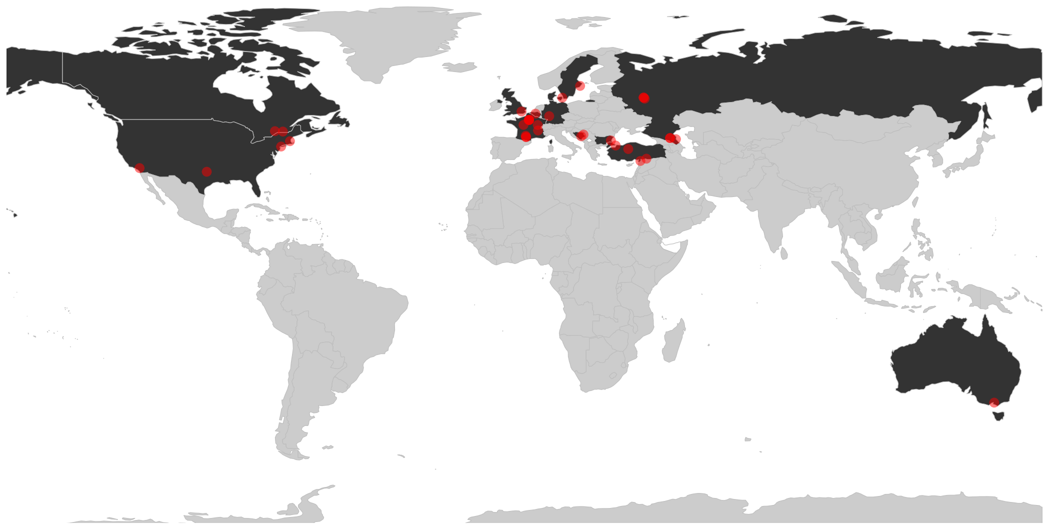
The goal of this study is to systematically measure changes in anti-Muslim hostility following unexpected terrorist attacks in Western countries. In this section, I describe how I generated an original, tweet-level dataset of anti-Muslim hostility, which geographically attributes Twitter posts to regions affected by terrorist attacks.

[6] Studies have examined the effects of other attacks, including the July 7, 2005 attacks in London and the March 11, 2004 attacks in Madrid (Fischer et al., 2007; Hanes and Machin, 2014; Montalvo, 2011). However, to the best of my knowledge, no study has examined the impact of multiple attacks over several years.

3.3.1 Terrorist attacks

As a first step, I collected data from news reports on the timing and location of acts of terror perpetrated by radical jihadists in Western countries from 2010 to 2016. These attacks include, among others, the 2015 Paris attacks in France, the 2016 Brussels bombing in Belgium, and the 2016 Orlando nightclub shootings in the United States. Figure 3.2 displays these attacks visually. Combined, these acts of violence killed almost 800 individuals and injured over 2,000. Table 3.11 in the Appendix provides additional details on these attacks.

Figure 3.2: Terrorist Attacks by Radical jihadists in the West (2010 to 2016)



Note: The figure displays the locations of terrorist attacks perpetrated in Western countries by individuals identifying with radical jihadi ideology from 2010 to 2016. Information on thee timing and location of attacks was obtained by the author from news reports.

As I discuss further in section 3.4, the locations of terrorist attacks are non-random and likely chosen specifically to achieve the organization's strategic goals. It is thus impossible to compare anti-Muslim content cross-sectionally between targeted and non-targeted locations and ascribe a causal interpretation to observed differences. Indeed, locations may be chosen for terrorist attacks precisely because they have high levels of anti-Muslim hostility! For this reason, in this study, my identification strategy relies on high-frequency comparisons of anti-Muslim content *within* locations, attacks, and individuals. Terrorist attacks are very likely unexpected and thus "(i.e., random) in time," facilitating a causal interpretation of a discontinuous jump in anti-Muslim content immediately prior to and following the terrorist attack.

3.3.2 The targeted population

To measure anti-Muslim content generated by the targeted population before and after attacks, I used the Twitter Application Program Interfaces (APIs) to collect data.[7] It is technically challenging to obtain an unbiased estimate of tweet-level data in a particular country before and after attacks. One reason is that Twitter users selectively enable geo-location for tweets and user accounts, and relying on this sparse information is less than ideal. (Unlike in the first chapter of this dissertation, I did not have a network of accounts of interest from which to build lists of followers and use geolocation prediction.) Ensuring that a sample of users is sufficiently unbiased is thus a non-trivial challenge. As a solution, I employed Twitter's Streaming APIs[8] to obtain a random sample of Twitter users in the *time zone* of the targeted country. To be sure, this is a noisy measure, but it the most likely to yield an unbiased measure of sample inclusion.

For each attack, I randomly sampled about a thousand users who generated content in the time zone of the attack.[9] I then traversed their historical timelines and downloaded the entirety of their Twitter content in the month prior to and following the attack. To be sure, this sample is not necessarily likely to be representative of the entirety of the Twitter population — because it excludes those who were not tweeting at the time of data collection

[7] `https://dev.twitter.com/overview/documentation`

[8] `https://dev.twitter.com/streaming/overview`

[9] For a small number of attacks, there was insufficient Twitter data to create a user-level panel around the event. These attacks are the July 18, 2012 Burgas bus bombing in Bulgaria and the October 5, 2015 bombings in Grozny, Russia.

or who have created new accounts that lacked tweets in the weeks surrounding the attack — but there is no reason to believe that these imperfections in measurement are likely to bias a high-frequency comparison of content generated by those accounts that were open at the time. Nonetheless, any attempt to generalize the results of this study beyond those accounts which were presently tweeting and also tweeting at the time of the attack should be viewed with caution.

From an empirical design standpoint, I also sought to ensure that I would have a sufficient number of Twitter posts to facilitate a within-user comparison of anti-Muslim content before and after the attack. As discussed further below, a within-user comparison is necessary to control for unobserved heterogeneity between users that may lead to the incorrect conclusion that the attack had a causal effect on anti-Muslim sentiment, where in fact any observed difference may simply be compositional in nature, reflecting a different "mix" of users before and after the attack. I thus applied a series of filters to ensure that I could use a panel design and perform within-user comparisons of changes in anti-Muslim content before and after each attack.

I began with 124,619 users and 5,816,565 tweets in my dataset. The first filter that I applied was technical in nature: I simply removed those users whose account was created after the attack, as it is impossible to calculate a within-user effect when the account was created after the attack. The second filter I applied was to remove users with more than 3,200 tweets after the attack. This filter was used because Twitter limits the retrieval of tweets to the most recent 3,200. If a user had more than 3,200 tweets since the attack, then that all of his/her available tweets were generated after the attack, making it impossible to measure their pre-attack content. Finally, I removed those users with less than 20 tweets in the month before and less than 20 tweets in the month after the attack. Without this filter, there would be an insufficient number of tweets in the pre-post periods to facilitate an empirically meaningful within-user study. After applying these filters, I have 116,942 users in my dataset and 4,991,795 tweets, which constitute about 94% and 86%, respectively, of the original unfiltered sample.

Creating measures of anti-Muslim content

After obtaining a sample of individuals meeting the inclusion criteria described above, I downloaded the full historical timelines of these users using the Twitter REST APIs.[10] As described in the Twitter API documentation and in Chapter 2 of this dissertation, there are many data fields returned at the tweet-level, including: the UTC timestamp when the tweet was created; the geographical coordinates of the tweet, if supplied by the user; the language of the tweet; whether it was a 'retweet' or quoting another tweet, and if so, the identification information of that tweet; and the raw text of the tweet. Unlike in Chapter 2, where the network structure of "followers" and "friends" was crucial to the analysis, this chapter focuses primarily on anti-Muslim sentiment (textual content), and thus I disregard fields other than the textual content of the tweet, the user to which it belongs, and the exact timestamp of the creation of the tweet.

In order to generate an empirical measure of anti-Muslim hostility, it was necessary to obtain a "reference vocabulary" of words most likely, in a probabilistic sense, to reflect anti-Muslim sentiment. I employed two sources of data to generate this sort of vocabulary. First, I queried the REST API for Twitter posts[11] that included common anti-Muslim hashtags described in Table 3.1. These sorts of hashtags are often utilized by users to "label" the nature of the tweet in an abbreviated form. For this reason, the text contained in tweets with these hashtags is likely to reflect the nature of the hashtags themselves, i.e., anti-Muslim sentiment. To be sure, occasionally tweets labeled with these hashtags may contain the *opposite* sentiment, i.e., a sarcastic critique of an anti-Muslim position. But

[10] `https://dev.twitter.com/rest/public`

[11] $N = 288,886$

those are relatively infrequent compared to content consistent with the hashtag; thus, the hashtag selection procedure is unlikely to yield systematically biased content, even if it is noisy some of the time. A sample of these hashtag-labeled tweets is shown in Table 3.2.

Table 3.1: Anti-Muslim hashtags on Twitter

#banislam	#banmuslims	#BanShariaLaw	#IslamIsTheProblem	#hatemuslims
#IslamIsEvil	#NoMuslimRefugees	#eradicateislam	#islamstupidity	#ExterminateIslam
#BanSharia	#StopRapeJihad	#BanTheBurka	#noislam	
#hateislam	#StopIslam	#islamoutnow	#NoSharia	

Table 3.2: Example of Anti-Muslim tweets

Don't get me wrong: I don't want all the muslims gone. I just want them not to be muslims. The religion is evil. #EradicateIslam
offending muslims is a moral obligation if you dont then you destroy your freedom and future. #eradicateislam
CHRISTIANITY under ATTACK! Ever wonder why moslems literally get away W/ murder & ZERO CONSEQUENCES? #BanIslam
Why #Islam is not compatible with the #West ...#banislam in the West
there's no assimilation by moslems they create no-go sharia jones in every country that welcomes them. #banislam
You can't #coexist with people who want to impoverish and kill you ... #IslamIsTheProblem #StopIslam
Remember it is your RIGHT to refuse service from a Moslem Dr. This isn't racism, it's survival. #banMuslims#banIslam#WakeUpAmerica

As a second source for a "reference vocabulary," I used all tweets posted by politicians from far-right parties who oppose Muslims in Europe.[12] As discussed in detail in Chapter 2, contemporary European far-right parties play a central role in spreading anti-Muslim sentiment across Europe. While their platforms cross many social issues, most parties advance a populist and exclusionary agenda targeting Muslims and other minorities (Golder, 2016). As many politicians use Twitter to advance their agenda and gain supporters (Grant, Moon and Busby Grant, 2010; Theocharis et al., 2016), I used information on the tweets posted by about 90 politicians from far-right parties.

[12] $N = 196,944$

Specifically, I manually collected a list of far-right politicians and their Twitter handles, which is described in Table 3.10 in the Appendix. I then proceeded to systematically download the timelines of these Twitter accounts, including the text of each tweet found on their timelines. Table 3.3 shows examples of these tweets. While not all of the tweets in these far-right accounts are specifically directed towards Muslims (i.e., some refer to immigration or far-right views more generally), these examples nonetheless suggest that many do reflect anti-Muslim sentiment in particular, and thus the textual content found in these timelines can serve as a meaningful and useful outcome for the empirical analysis in this project.

Table 3.3: Examples of anti-Muslim tweets by far-right politicians

Party	**Politician**	**Handle**	**Tweet**
Liberty GB	Paul Weston	@paulwestonlibgb	1/3 of all muslims want to stone women to death for adultery and murder people for leaving islam. How peaceful is that?
Liberty GB	Jack Buckby	@jackbuckby	this weird anomaly whereby muslims =4% of the population but 90% of paedophile rape gang convictions. strange.
UKIP	Bill Etheridge	@BillDudleyNorth	UKIP MEP says locals horrified Labour Council are taking bribe to build a mega mosque. Labour pro muslim anti British
Vlaams Belang		@VlaamsBelangBru	"Not all muslims are terrorists" is like an airline with the world's worst fatality rate saying "Not all of our planes crash"... #Islam
Vlaams Belang	Anke Van dermeersch	@Anke_online	A consideration about the relation between the West and islam: tolerance becomes a crime when applied to evil
Vlaams Belang	Filip Dewinter	@FDW_VB	@realDonaldTrump is right! Close the borders for muslims! 'Obama, you let'me in, I'll send ALL of them back.'

I applied standard text pre-processing procedures to the raw text of these tweets: removal of numbers, punctuation and stopwords, word stemming and vectorization, which resulted in document-term matrices for each source. A document-term matrix is a $n \times k$ matrix of word frequencies, where each document is a row and each term in the vocabulary of these documents is a column. Element i, j of this matrix is the frequency of word $j \in 1, ..., k$ in document $i \in 1, ..., n$. An example of a document-term matrix for the sentences "The cat jumped over the moon." and "The dog jumped over the cat." is shown in Table 3.4:

Table 3.4: Example of document-term matrix

	the	cat	dog	jumped	over	moon
(Sentence 1)	2	1	0	1	1	1
(Sentence 2)	2	1	1	1	1	0

I generated document-term matrices for the anti-Muslim sentiment and far-right content. These term matrices represent the benchmark vocabularies to which I compared the content of each tweet generated by individuals in countries targeted by terrorism. As noted previously, these document-term matrices count the number of times that words in the composite vocabulary (i.e., the union of words found in all of these sources) appeared in the tweet.

I calculated anti-Muslim and far-right similarity scores for each tweet as the cosine similarity between the words in the tweet and the anti-Muslim and far-right vocabularies. The cosine similarity between two document-term vectors **a** and **b**, each of length $k \times 1$, is calculated as follows:

$$similarity = \frac{\sum_{j=1}^{k} a_j b_j}{\sqrt{\sum_{j=1}^{k} a_j^2}\sqrt{\sum_{j=1}^{k} b_j^2}}$$

where a_j is the number of times term j appears in document **a** and b_j is the number of times term j appears in document **b**. In the example given in Table 3.4, the cosine similarity is given by:

$$similarity = \frac{\sum_{j=1}^{k} a_j b_j}{\sqrt{\sum_{j=1}^{k} a_j^2}\sqrt{\sum_{j=1}^{k} b_j^2}} = \frac{7}{2.83 \cdot 2.83} \approx 0.875$$

In general, the similarity score ranges from 0 (not similar) to 1 (completely similar). In this study, I derived the cosine similarity between each user's tweet and each of the two anti-Muslim vocabularies (i.e., the hashtag-based and far-right party-based) in their entirety. For ease of interpretation, I standardized the similarity scores in the statistical analysis.

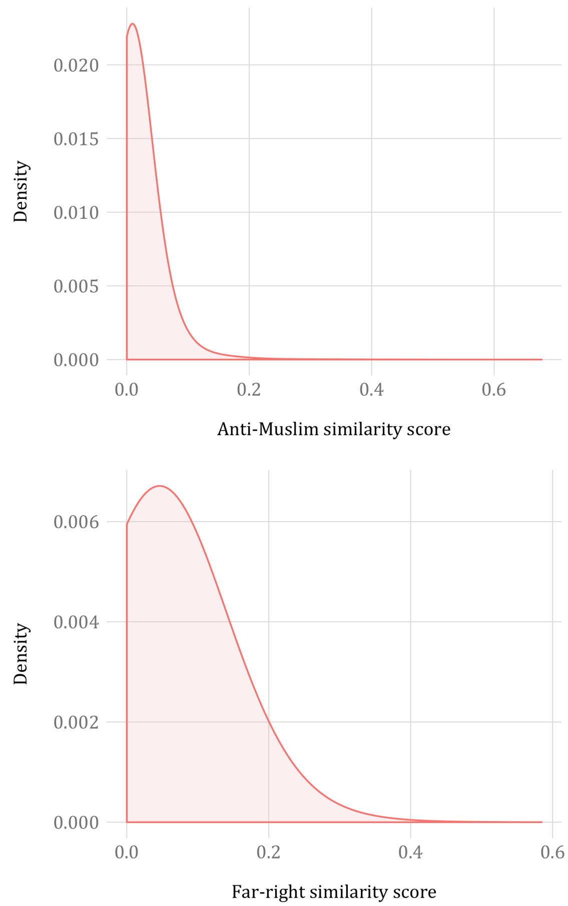
I created a tweet-level dataset for each user in the sample, which included tweet-level covariates (e.g., the timestap and the text of the tweet), user-level covariates (e.g., name, profile description, number of followers, country), and information related to the attack to which the individual was exposed (e.g., location and timing of the attack, number of killed and injured). I then combined the individual-level datasets to a master all-user dataset for each attack. Then, I merged 36 datasets for each attack into one pooled all-attacks dataset. This enabled me to compare the level of anti-Muslim posting before and after terrorist attacks for thousands of individuals in targeted countries. Summary statistics for this dataset are given in the Table 3.5. Figure 3.3 shows density plots for the anti-Muslim and far-right similarity scores. It can be seen that these outcome measures are skewed, which means that interpretation of standard deviation units should take into account that the standard deviation of these outcome variables is larger than would be the case with a normally distributed variable. This would render coefficients smaller.

Table 3.5: Summary statistics

Statistic	N	Mean	St. Dev.	Min	Max
After attack = 1	4,991,794	0.572	0.495	0	1
Anti-Muslim similarity score	3,250,107	0.017	0.031	0	0.678
Far-right similarity score	3,119,581	0.054	0.051	0	0.585
Number killed in terrorist attack	4,991,795	52.687	59.095	0	224
Number injured in terrorist attack	4,815,953	136.976	159.182	0	400
Number of attacks in country	4,991,795	6.424	3.112	1	10

Note: The table reports summary statistics for the variables used in the analysis. The unit of analysis is the tweet.

Figure 3.3: Similarity scores



Note: The figure displays density plots for the anti-Muslim and far-right similarity scores for all tweets in the dataset.

3.4 Empirical strategy

One of the main challenges with studying the effect of exposure to violence, especially terrorism, is that such type of violence is not random in many cases. Perpetrators of terrorism strategically choose the location and general timing of their attacks in order to increase the number of victims or publicize their cause (Crenshaw, 1981; Kydd and Walter, 2006). As a result, targeted populations, especially in locations undergoing conflict, may anticipate attacks and change their behavior accordingly. Indeed, in general, a comparison between targeted and non-targeted populations would be confounded by the very differences that gave rise to the targeting in the first place. Any sort of cross-sectional comparison would undoubtedly be biased.

I seek to mitigate this challenge in two ways. First, I study contexts in which violence is not ongoing. Western countries experiencing terrorism in recent years were not fighting civil wars or insurgencies in their territories at the same time. Hence, residents of these countries were likely living in an ordinary fashion, not expecting to be targeted by terror. The sudden, unexpected exposure to a terrorist attack of individuals who are not generally exposed to ongoing violence thus implies that behavior immediately prior to the attack can more credibly serve as a counterfactual for behavior following the attack.

Second, I use high-frequency data on Twitter behavior to measure changes in Tweet content immediately before and after terrorist attacks, exploiting the hour/minute/second timestamps of individual tweets. The identifying assumption is that while the general timing of terrorist violence may be anticipated in the population (as a result of government security agencies' counter-terrorism warnings, for example), the exact moment in time in which the attack is performed is likely to be unexpected. For that reason, the comparison of online rhetoric before and immediately after a terrorist attack facilitates causally attributing any observed changes to the attack. Moreover, I use a large number of attacks to strengthen the causal interpretation of these findings. While other, confounding events spuriously coinciding with one attack might be driving the estimated effect for that attack, these elements are unlikely to be present across many attacks taking place in different points in time across various countries. Furthermore, I complement this analysis with placebo tests using fictional attack times.

3.4.1 Primary model

As noted previously, in this study I seek to estimate the difference between individuals' Twitter posts before and after terrorist attacks, for thousands of individuals in over thirty attacks in twelve countries. My primary estimations employ the following least squares regression model:

$$y_{p,i,j,k} = \beta T_{p,i,j,k} + \alpha_i + \gamma_j + \delta_k + \varepsilon_{i,j,k} \tag{3.1}$$

where $y_{p,i,j,k}$ is the cosine similarity score to the anti-Muslim (or far-right vocabulary) of Twitter post p generated by user i before or after attack j in country k; and where T is an indicator coded 0 for tweets generated before a terrorist attack and 1 afterwards. All models include user, attack, and country fixed effects (α_i, γ_j, and δ_k respectively), to account for unobserved, time-invariant heterogeneity between countries, attacks, and individual users. Standard errors are clustered at the user level to account for serial correlation in tweet content posted by the same user.

As noted previously, the causal interpretation of the β coefficient is substantially strengthened by the use of high-frequency data. The presence of hour/minute/second timestamps facilitates comparing tweet content immediately before and immediately after each terrorist attack, whose precise timing is likely to be highly unexpected. Thus, my primary estimations limit the subsample to tweet content generated one day before and after each terrorist attack, to minimize the possibility that additional content arriving further away from the terrorit attack itself might spuriously bias this estimate. However, I also examine additional time windows of greater length — specifically, +/- 7 days and +/- 30 days — to evaluate how sensitive the results are to the inclusion of additional data further away from the event.

3.4.2 Additional specifications

I also present results from additional estimations. First, I examine the time horizon of the effect, i.e., the extent to which the effect dissipates with time. It is possible that anti-Muslim content may sharply rise immediately after an attack, but then moderate as the harsh memories of the attack fade away. I thus re-estimate the β coefficient with varying "post" windows, holding fixed the "pre" window at -7 days, and considering how the β coefficient changes as the "post" window is extended each day.

Second, I consider whether the β coefficient varies heterogeneously with the number of casualties of each terrorist attack. To generate these data, I manually combed through news reports about each attack and recorded the number of individuals killed (or injured) in the attack. I estimate the following interaction specification:

$$y_{p,i,j,k} = \beta_1 T_{p,i,j,k} + \beta_2 N_{j,k} + \beta_3 (T_{p,i,j,k} \times N_{j,k}) + \alpha_i + \gamma_j + \delta_k + \varepsilon_{i,j,k} \tag{3.2}$$

where $y_{p,i,j,k}$ is the anti-Muslim sentiment (i.e., cosine similarity score) of Twitter post p generated by user i before or after attack j in country k; $N_{j,k}$ is the number of individuals killed (or injured) in attack j in country k; and T is an indicator coded 0 for tweets generated before a terrorist attack and 1 afterwards. In this specification, β_3 is the primary coefficient of interest, as it reflects the extent to which the change in anti-Muslim sentiment varies with the number of casualties. As with the primary estimation, all models include user, attack, and country fixed effects to account for unobserved, time-invariant heterogeneity between countries, attacks, and individual users. Standard errors are clustered at the user level to account for serial correlation, as before.

Finally, I estimate a placebo test which replicates the primary analysis, replacing the terrorist attack dates with randomly selected dates. Specifically, each placebo date was set to be one week before each attack took place. A statistically significant β coefficient on these arbitrary dates would undermine the causal interpretation of the results, as it may indicate a heavy-tailed distribution of anti-Muslim content (i.e., such that the random selection of any two dates would result in a significant difference under the standard t- or normal distribution, even if the underlying effect was null).

3.5 Results

I begin with the primary specification. As Table 3.6 shows, there is a modest but statistically significant increase in anti-Muslim (and far-right) content following terrorist attacks, on the order of magnitude of .06 standard deviations for the anti-Muslim measure, and .11 standard deviations for the far-right measure when employing a window length of [-1, +1] days surrounding attacks. Moreover, the effect remains statistically significant as longer window lengths are employed. While this effect seems small in magnitude, it is worth noting that this sort of textual comparison is very noisy: tweets can address a wide range of topics, and word frequencies are an imperfect measure for underlying content. For this reason, it is difficult to make a strong quantitative inferences about the magnitude of the change from the difference in cosine similarity measure. But we can conclude with statistical confidence that the proportion of anti-Muslim and far-right content systematically increased in the wake of these unexpected terrorist attacks, even when using a very short window of time around each attack. Recall, these results reflect changes *within users*, showing that individuals in countries targeted by terrorism posted more tweets reflecting anti-Muslim sentiment after attacks compared to what they posted before the attacks. These results systematically hold across thirty-six attacks in twelve countries.

Figures 3.4 and 3.5 examine how the difference in anti-Muslim and far-right content varies with the length of the "post" window employed. It can be seen that the estimated coefficient decreases smoothly as the time window increases, but remains significantly different from zero. This suggests that the immediate effect of a terrorist attack is to produce a

Table 3.6: Terrorist attacks and anti-Muslim and far-right rhetoric

Anti-Muslim content (sd units)			
	[-1,1]	[-7,7]	[-30,30]
After attack = 1	0.060***	0.029***	0.014***
	(0.006)	(0.003)	(0.002)
User fixed effects	✓	✓	✓
Attack fixed effects	✓	✓	✓
Country fixed effects	✓	✓	✓
R^2	0.183	0.147	0.144
Number of clusters	10,637	12,850	13,979
Number of observations	229,740	970,522	3,249,622
Far-right content (sd units)			
	[-1,1]	[-7,7]	[-30,30]
After attack = 1	0.110***	0.051***	0.020***
	(0.007)	(0.003)	(0.002)
User fixed effects	✓	✓	✓
Attack fixed effects	✓	✓	✓
Country fixed effects	✓	✓	✓
R^2	0.094	0.064	0.053
Number of clusters	9,888	12,128	13,484
Number of observations	225,752	935,499	3,118,955

Standard errors in parentheses, clustered at the user level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

sharp spike in anti-Muslim (and far-right) sentiment among targeted populations, followed by a partial reversion to a still-elevated level of hostility. Of course, this does not imply that the effect itself is short-lived: as discussed at length in Chapter 2 of this dissertation, anti-Muslim hostility is widespread in many Western countries; terrorist attacks are thus periods in which such animosity intensifies. Indeed, it is interesting that the effect appears to remain significant for an extended period of time, as one might have predicted merely only a transient spike in the wake of the attack. This pattern suggests that terrorist attacks may expose Muslim minorities to protracted periods of anti-Muslim sentiment in countries targeted by terrorism, furthering the strategic recruitment goals of armed organizations.

Next, Tables 3.7 and 3.8 consider whether the increase in anti-Muslim and far-right sentiment differs by the number of casualties in each attack. The tables show that the

Figure 3.4: The impact of terrorist attack on anti-Muslim rhetoric over time

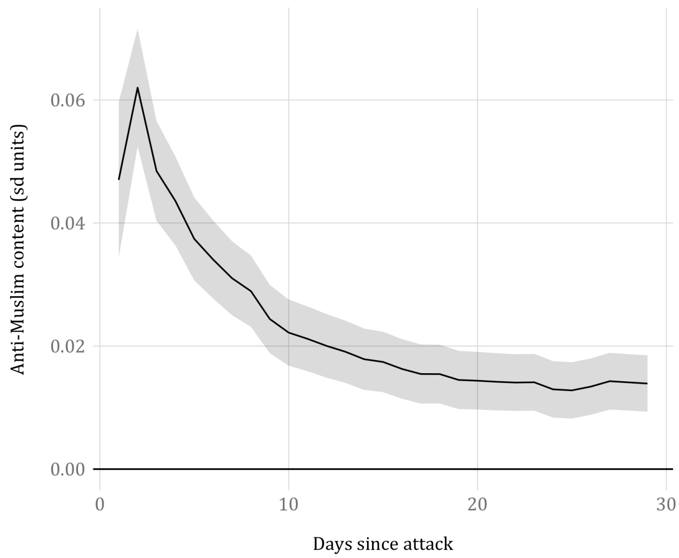
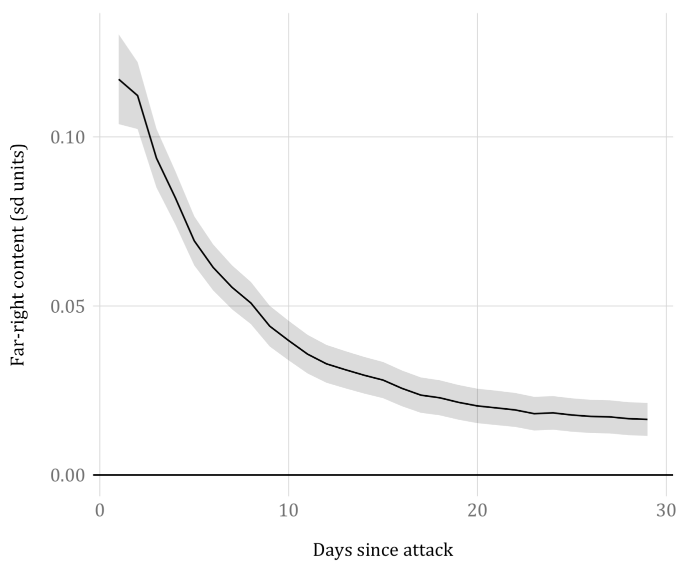


Figure 3.5: The impact of terrorist attack on far-right rhetoric over time



interaction term is positive in all specifications and statistically significant in all but two, indicating the effect increases with the number of people killed and injured. The insignificant point estimates are similar in magnitude, suggesting that the lack of significance may be due to insufficient power at the [-1,+1] window.

For the anti-Muslim outcome (Table 3.7), the results indicate that a one standard deviation increase in the number of casualties is associated with an increase of 0.004 to 0.005 standard deviations in anti-Muslim content, depending on the specification employed. This is a relative increase of 18% to 28% over the baseline coefficient of 0.014 to 0.028 standard deviations, which is the pre-post increase in anti-Muslim content at the mean number of casualties.[13] Similarly, a one standard deviation increase in the number of people injured is associated with an increase of 0.004 to 0.007 standard deviations in anti-Muslim content, a relative increase of 24% to 28% over the baseline coefficient of 0.014 to 0.029 standard deviations.

For the far-right outcome (Table 3.8), the results show that a one standard deviation increase in the number of casualties is associated with an increase of 0.019 to 0.059 standard deviations in far-right content, depending on the specification employed. This is a relative increase of 58% to 100% over the baseline coefficient of 0.02 to 0.101 standard deviations. Similarly, a one standard deviation increase in the number injured is associated with an increase of 0.016 to 0.051 standard deviations in far-right content, a relative increase of 52% to 84% over the baseline coefficient of 0.019 to 0.097 standard deviations. To sum up, the increase in anti-Muslim and far-right content after terrorist attacks is greater for attacks with more casualties, though this finding is stronger and more consistent in terms of statistical significance for the far-right outcome than the anti-Muslim outcome. These are important results, as they suggest that targeted populations increase hostile rhetoric towards Muslims after more lethal terrorist acts. Radical jihadi groups seeking to provoke targeted populations to become hostile to Muslims might seek to perpetrate more deadly attacks to increase the chances of alienation and recruitment.

[13] As the number of casualties is standardized, the "After attack = 1" coefficient is the mean increase in anti-Muslim content at the mean number of casualties.

Table 3.7: Terrorist attacks and anti-Muslim rhetoric, by number of victims

	Anti-Muslim content (sd units)		
Days before and after attack	[-1,1]	[-7,7]	[-30,30]
	Number killed		
After attack = 1	0.060***	0.028***	0.014***
	(0.006)	(0.003)	(0.002)
Number killed (sd units)	0.001	0.008	0.016**
	(0.021)	(0.008)	(0.008)
After attack = 1× Number killed (sd units)	0.002	0.005**	0.004**
	(0.006)	(0.003)	(0.002)
User fixed effects	✓	✓	✓
R^2	0.183	0.147	0.144
Number of clusters	10,637	12,850	13,979
Number of observations	229,740	970,522	3,249,622
	Number injured		
After attack = 1	0.062***	0.029***	0.014***
	(0.007)	(0.003)	(0.002)
Number injured (sd units)	0.056	0.040*	0.044*
	(0.040)	(0.023)	(0.023)
After attack = 1× Number injured (sd units)	0.005	0.007**	0.004**
	(0.006)	(0.003)	(0.002)
User fixed effects	✓	✓	✓
R^2	0.177	0.141	0.138
Number of clusters	10,107	12,236	13,347
Number of observations	221,986	931,676	3,108,770

Standard errors in parentheses, clustered at the user level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

An interesting question arising from these findings is whether the results are driven by individuals who already displayed hostility to Muslims before the attack or by individuals who have been benign in their rhetoric before being exposed to terrorism. Figure 3.6 shows density plots of the estimated effects of terrorist attacks on anti-Muslim and far-right content for individuals who expressed high levels hostility before the attack (red) and for those expressing low levels of hostility (blue). For both outcomes, the difference in rhetoric is almost identical for the two groups, suggesting that terrorism leads individuals in targeted countries to express greater hostility to Muslims, regardless of whether they were already hostile before the attacks. This is an important finding, indicating that terrorism indiscriminately provokes targeted populations to increase anti-Muslim hostility.

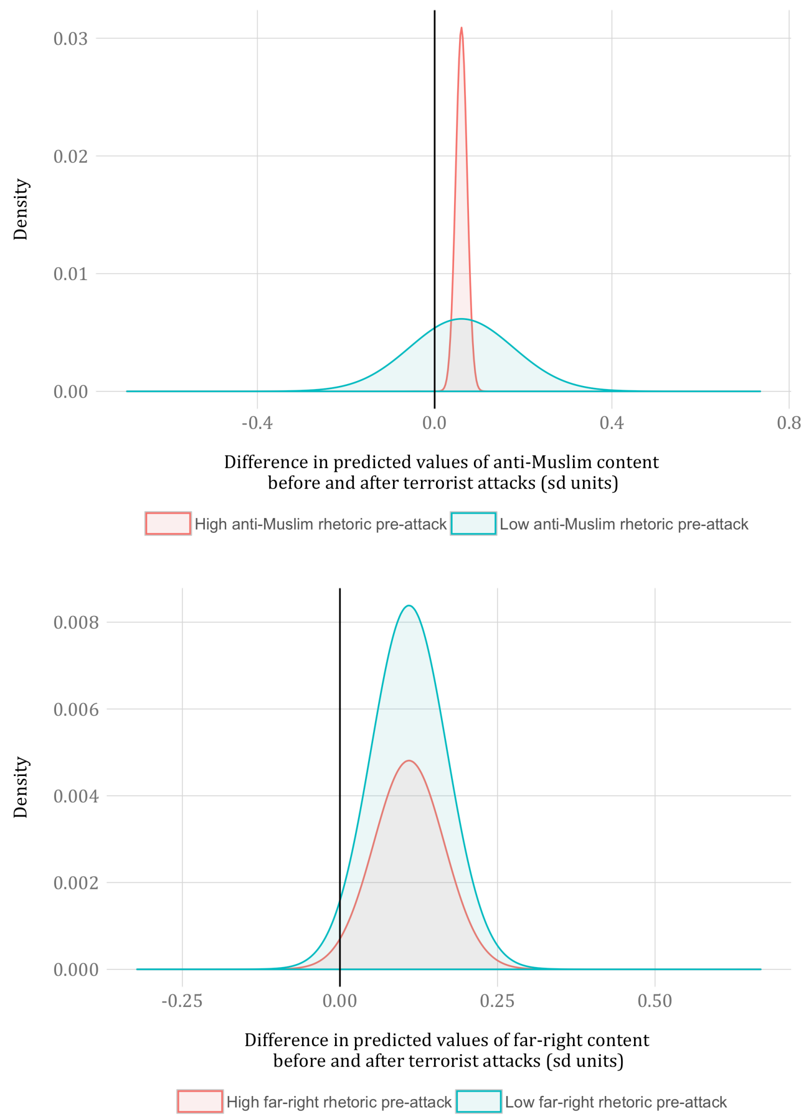
Table 3.8: Terrorist attacks and far-right rhetoric, by number of victims

	Far-right content (sd units)		
Days before and after attack	[-1,1]	[-7,7]	[-30,30]
	Number killed		
After attack = 1	0.101***	0.048***	0.020***
	(0.007)	(0.003)	(0.002)
Number killed (sd units)	-0.047	-0.058***	-0.010
	(0.051)	(0.018)	(0.012)
After attack = 1× Number killed (sd units)	0.059***	0.048***	0.019***
	(0.007)	(0.003)	(0.002)
User fixed effects	✓	✓	✓
R^2	0.095	0.065	0.053
Number of clusters	9,888	12,128	13,484
Number of observations	225,753	935,499	3,118,955
	Number injured		
After attack = 1	0.097***	0.044***	0.019***
	(0.006)	(0.003)	(0.002)
Number injured (sd units)	-0.035	-0.045**	-0.007
	(0.042)	(0.018)	(0.011)
After attack = 1× Number injured (sd units)	0.051***	0.041***	0.016***
	(0.006)	(0.003)	(0.002)
User fixed effects	✓	✓	✓
R^2	0.094	0.064	0.053
Number of clusters	9,782	11,868	13,083
Number of observations	224,932	931,263	3,102,422

Standard errors in parentheses, clustered at the user level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Finally, in Table 3.9, I present the results from a placebo test for the [-1,+1] and [-7,+7] windows, in which I analyze the 'effect' of arbitrary days on in anti-Muslim and far-right content. For each attack, the placebo date was set to be one week prior to the actual date of the attack. The results show that none of the coefficients are statistically significant in the placebo test, and their signs and magnitude are inconsistent across window lengths. This test provides additional evidence that the reported results in this chapter are unlikely to have occurred by random chance.

Figure 3.6: The effect of terrorist attacks on anti-Muslim and far-right rhetoric, by baseline rhetoric before attacks



Note: The figure shows density plots of the differences in anti-Muslim and far-right rhetoric after terrorist attacks, for individuals who already posted high levels of such content before the attacks (marked in red) and for individuals who posted low levels of such context before the attacks (marked in blue). It can be seen that increased anti-Muslim and far-right rhetoric after terrorist attacks change in the same way for both groups.

Table 3.9: Placebo tests

	Anti-Muslim content		Far-right content	
Days before and after placebo date	[-1,1]	[-7,7]	[-1,1]	[-7,7]
After placebo date = 1	0.032	-0.005	0.060	0.004
	(0.033)	(0.010)	(0.043)	(0.010)
User fixed effects	✓	✓	✓	✓
Attack fixed effects	✓	✓	✓	✓
Country fixed effects	✓	✓	✓	✓
R^2	0.211	0.165	0.109	0.064
Adjusted R^2	0.158	0.151	0.051	0.049
Number of clusters	9,673	12,634	8,993	11,926
Number of observations	155,092	797,892	149,037	762,867

Standard errors in parentheses, clustered at the user level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.6 Conclusion

This chapter has sought to examine whether terrorist attacks perpetrated by jihadi extremists in the West provoke targeted populations to become hostile to the crowd from which jihadi groups seek to recruit supporters. Unlike prior work on the provocation strategy of terrorism that has focused almost exclusively on the behavior of targeted governments (Bueno de Mesquita and Dickson, 2007; Lake, 2002; Rosendorff and Sandler, 2010), this chapter showed that terrorist attacks systematically affect the behavior of individuals in targeted countries. Specifically, the findings show that acts of terror carried out by radical jihadists in the West significantly increase anti-Muslim and far-right rhetoric among tens of thousands of people in targeted countries.

This pattern is important, especially in light of the findings presented in Chapter 2, which show that there is a strong link between measures of anti-Muslims hostility and far-right popularity and support for Islamic State in the West. Groups like ISIS capitalize on anti-Muslim hostility to recruit supporters, and might be using terrorist attacks to manipulate levels of animosity precisely for recruitment purposes. In a recent interview, a jihadist ISIS fighter stated that “We don’t need to convince Muslims in the Middle East that the West is against them. They already know. The next step for the Islamic State is to reach Muslims in America and Europe" (Revkin and Mhidi, 2016). Terrorist attacks that increase hostility towards Muslims serve as a tool for groups like ISIS to persuade potential supporters that the West is against them. Such message might be powerful for individuals facing indiscriminate hostility by fellow citizens, especially as it intensifies in the wake of terrorist acts.

The results also contribute to literature on the psychology of exposure to terrorism. Earlier in this chapter, I showed that attacks increase hostility against Muslims in a similar manner for individuals who already express high levels of anti-Muslim hostility before the attacks as well as for those who do not. This suggests that there is something inherent about acts of terror that affects targeted citizens with different prior behaviors in the same way. Theories of inter-group hostility might argue that this factor is the fear and threat that attacks generate, which can lead people, regardless of their prior beliefs, to abandon commitments to liberal values, tolerance, and inclusion (Marcus, 1995; Sullivan, Piereson and Marcus, 1993). In addition, the finding show that anti-Muslim hostility increases more strongly after attacks with more victims, indicating that the animosity might be driven by people's sensitivity casualties.

Finally, this chapter speaks to the vicious cycle of radicalization and hostility presented in Chapter 2. In the vicious cycle, anti-Muslim hostility drives pro-ISIS radicalization, but visible forms of extremism, such as terrorist attacks, also lead to greater levels of hostility. This chapter shows this pattern empirically by demonstrating that terror attacks carried out by individuals who radicalized to support jihadi ideology significantly increase anti-Muslim hostility in the West. This mutual feedback loop is something that many contemporary Western societies encounter, and requires planning of possible solutions. How can societies break the cycle of radicalization and hostility? In the next chapter, I evaluate one solution to radicalization in which Western governments have been investing in recent years.

3.7 Appendix

Table 3.10: Twitter handles of far-right politicians in Europe

Country	Party	Politician	Twitter handle
France	Front National		@FN_officiel
France	Front National	Jean-Marie Le Pen	@lepenjm
France	Front National	Marine Le Pen	@MLP_officiel
France	Front National	Louis Aliot	@louis_aliot
France	Front National	Marie-Chris Arnautu	@MCArnautu
France	Front National	Jean-Francois Jalkh	@JFJalkh
France	Front National	Florian Philippot	@f_philippot
France	Front National	Steeve Briois	@SteeveBriois
France	Front National	France Jamet	@JametFrance
France	Front National	Dominique Bilde	@DominiqueBilde
France	Front National	Frederic Boccaletti	@FnVar_officiel
France	Front National	Gilbert Collard	@GilbertCollard
France	Front National	Bruno Gollnisch	@brunogollnisch
France	Front National	Michel Guiniot	@MichelGuiniot
France	Front National	Gilles Lebreton	@Gilles_Lebreton
France	Front National	Marion Le Pen	@Marion_M_Le_Pen
France	Front National	Dominique Martin	@DMartinFN
France	Front National	Joelle Melin	@JoelleMelinFN
France	Front National	Bernard Monot	@Bernard_Monot
France	Front National	Sophie Montel	@Sophie_Montel
France	Front National	Mireille d'Ornano	@MireilledOrnano
France	Front National	David Rachline	@david_rachline
France	Front National	Thibaut delaTocnaye	@TdlTocnaye
Germany	NPD		@npdde
Germany	NPD		@npdthueringen
Germany	NPD	Frank Franz	@FrankFranz
Germany	NPD	Ronny Zasowk	@RonnyZasowk
Germany	NPD	Sebastian Schmidtke	@SebastianNPD
Germany	NPD	Jens Baur	@Jens_Baur
Germany	NPD	Claus Cremer	@claus_cremer
Germany	NPD	Jean-C. Fiedler	@jean_fiedler
Germany	AfD		@AfD_Bund
Germany	AfD	Frauke Petry	@FraukePetry
Germany	AfD	Der Meuthen	@JoergMeuthen
Germany	AfD	Beatrix von Storch	@Beatrix_vStorch
Germany	AfD	Julian Flak	@JulianFlak
Germany	AfD	Armin Paul Hampel	@ArminPaulHampel
Germany	AfD	Georg Pazderski	@Georg_Pazderski
Germany	AfD	Andre Poggenburg	@PoggenburgAndre
Germany	AfD	Alice Weidel	@alice_weidel
UK	British Democrats		@BrDemocrats
UK	British Democrats	Andrew Brons	@andrewbronsmep
UK	British National Party		@bnp
UK	Liberty GB		@Liberty_GB
UK	Liberty GB	Paul Weston	@paulwestonlibgb
UK	Liberty GB	Jack Buckby	@jackbuckby
UK	Liberty GB	Theobald Wallingford	@TheoWallingford
UK	National Front		@NationalFrontTV

Twitter handles of far-right politicians in Europe (Cont.)

Country	**Party**	**Politician**	**Twitter handle**
UK	UKIP		@UKIP
UK	UKIP	Steven Woolfe	@Steven_Woolfe
UK	UKIP	Nigel Farage	@Nigel_Farage
UK	UKIP	Roger Helmer	@RogerHelmerMEP
UK	UKIP	Gerard Batten	@GerardBattenMEP
UK	UKIP	Paul Nuttall	@paulnuttallukip
UK	UKIP	Nathan Gill	@NathanGillMEP
UK	UKIP	Margot Parker	@MargotLJParker
UK	UKIP	Julia Reid	@julia_reid
UK	UKIP	Jane Collins	@Jane_CollinsMEP
UK	UKIP	David Coburn	@DavidCoburnUKip
UK	UKIP	Jonathan Arnott	@JonathanArnott
UK	UKIP	Patrick O'Flynn	@oflynnmep
UK	UKIP	Tim Aker	@Tim_Aker
UK	UKIP	Jill Seymour	@JSeymourUKIP
UK	UKIP	Jim Carver	@JamesJimCarver
UK	UKIP	Bill Etheridge	@BillDudleyNorth
UK	UKIP	Diane James	@DianeJamesMEP
UK	UKIP	John Bickley	@JohnBickleyUKIP
UK	UKIP	Peter Jewell	@Peter___Jewell
UK	UKIP	Douglas Carswell	@DouglasCarswell
UK	UKIP	Peter Whittle	@prwhittle
UK	UKIP	David Kurten	@davidkurten
UK	UKIP	Nathan Gill	@NathanGillMEP
UK	UKIP	Neil Hamilton	@NeilUKIP
UK	UKIP	Mark Reckless	@MarkReckless
UK	UKIP	David Rowlands	@DavidRowlandsAM
Belgium	Vlaams Belang		@vlbelang
Belgium	Vlaams Belang		@VlaamsBelangBru
Belgium	Vlaams Belang		@vbboom
Belgium	Vlaams Belang	Filip Dewinter	@FDW_VB
Belgium	Vlaams Belang	Bart Claes	@claesbart
Belgium	Vlaams Belang	Philip Claeys	@Philip_Claeys
Belgium	Vlaams Belang	Guy D'haeseleer	@GuydhaeseleerVB
Belgium	Vlaams Belang	Ortwin Depoortere	@OrtwinDepo
Belgium	Vlaams Belang	Chris Janssens	@chrisjanssensVB
Belgium	Vlaams Belang	Barbara Pas	@Barbara_Pas
Belgium	Vlaams Belang	Stefaan Sintobin	@StefaanSintobin
Belgium	Vlaams Belang	Anke Van dermeersch	@Anke_online
Belgium	Vlaams Belang	Tom Van Grieken	@tomvangrieken
Belgium	Vlaams Belang	Reccino Van Lommel	@reccino

Table 3.11: Terrorist Attacks by Radical jihadists in the West (2010 to 2016)

	Attack name	Date-time (GMT)	Location	Country	Killed	Injured
1	2010 Moscow Metro bombings	3/29/10 3:56	Lubyanka, Moscow	Russia	27	50
2	2010 Stockholm bombings	12/11/10 15:48	Stockholm	Sweden	0	2
3	Domodedovo Airport bombing	1/24/11 14:00	Domodedovo, Moscow	Russia	37	173
4	2011 Frankfurt Airport shooting	3/2/11 12:00	Frankfurt airport	Germany	2	2
5	Attack on US embassy in Bosnia	10/28/11 14:40	US Embassy, Sarajevo	Bosnia	0	1
6	Toulouse and Montauban shootings	3/11/12 15:00	Toulouse	France	0	1
7	Toulouse and Montauban shootings	3/15/12 14:00	Montauban	France	2	1
8	Toulouse and Montauban shootings	3/19/12 8:00	Toulouse	France	4	1
9	2012 Makhachkala attack	5/3/12 23:00	Makhachkala	Russia	14	87
10	Boston marathon bombings	4/15/13 18:49	Boston	USA	3	264
11	Reyhanli bombings	5/11/13 10:45	Reyhanli	Turkey	51	140
12	Murder of Lee Rigby	5/22/13 13:20	London	UK	1	0
13	2013 La Defense attack	5/25/13 16:00	Paris	France	0	1
14	Jewish Museum of Belgium shooting	5/24/14 13:50	Brussels	Belgium	4	0
15	2014 Endeavour Hills stabbings	11/23/14 8:45	Endeavour Hills, Victoria	Australia	0	1
16	2014 Grozny bombing	10/5/14 19:00	Grozny	Russia	5	12
17	2014 Saint Jean sur Richelieu attack	10/20/14 15:50	Saint-Jean-sur-Richelieu, Quebec	Canada	1	1
18	2014 shootings at Parliament Hill, Ottawa	10/22/14 13:52	Parliament Hill in Ottawa	Canada	1	3
19	Attack on police in NYC	10/23/14 18:00	Queens	USA	0	1
20	2014 Grozny clashes	12/4/14 22:00	Grozny	Russia	15	36
21	Stabbing in Joue-les-Tours	12/20/14 12:00	Joue-les-Tours	France	0	3
22	2014 Dijon attack	12/21/14 12:00	Dijon	France	0	11
23	Charlie Hebdo attack	1/7/15 10:30	Charlie Hebdo, Paris	France	12	11
24	Kosher supermarket attack	1/9/15 12:30	Porte de Vincennes, Paris	France	4	4
25	2015 Copenhagen shootings	2/14/15 14:00	Copenhagen	Denmark	2	5
26	Zvornik police station shooting	4/27/15 17:00	Zvornik	Bosnia and Herzegovina	1	2
27	Curtis Culwell Center attack	5/3/15 23:50	Garland, Texas	USA	0	1
28	Saint-Quentin-Fallavier attack	6/26/15 7:30	Saint-Quentin-Fallavier	France	1	11
29	2015 Suruc bombing	7/20/15 9:00	Suruc	Turkey	33	104
30	2015 Ankara bombings	10/10/15 7:04	Ankara	Turkey	102	400
31	Metrojet Flight 9268 attack	10/31/15 4:13	Sky	Russia	224	
32	November 2015 Paris attacks	11/13/15 20:20	Paris	France	130	368
33	2015 San Bernardino attack	12/2/15 18:59	San Bernadino	USA	14	22
34	2016 Istanbul bombing	1/12/16 8:20	Istanbul	Turkey	11	14
35	2016 Brussels bombings	3/22/16 7:00	Brussels	Belgium	31	300
36	2016 Orlando nightclub shooting	6/12/16 7:02	Orlando	USA	49	53

Table 3.12: Terrorist attacks and anti-Muslim and far-right rhetoric, by country

	(1) France	(2) Belgium	(3) United Kingdom	(4) United States	(5) Canada	(6) Turkey	(7) Russia
	Anti-Muslim content (sd units)						
After attack = 1	0.054***	0.070**	0.106**	0.090***	-0.009	0.110*	0.004
	(0.008)	(0.027)	(0.053)	(0.014)	(0.023)	(0.063)	(0.017)
User fixed effects	✓	✓	✓	✓	✓	✓	✓
Attack fixed effects	✓	✓	✓	✓	✓	✓	✓
R^2	0.109	0.177	0.144	0.131	0.149	0.363	0.173
Number of clusters	4,183	1,026	212	2,437	678	275	1,180
Number of observations	127,288	18,223	2,382	52,555	6,631	1,759	14,529
	Far-right content (sd units)						
After attack = 1	0.154***	0.071***	0.036	0.048***	0.021	0.101*	0.108**
	(0.009)	(0.025)	(0.049)	(0.011)	(0.024)	(0.061)	(0.055)
User fixed effects	✓	✓	✓	✓	✓	✓	✓
Attack fixed effects	✓	✓	✓	✓	✓	✓	✓
R^2	0.072	0.104	0.128	0.099	0.165	0.263	0.307
Number of clusters	4,196	1,202	211	2,426	678	292	220
Number of observations	127,896	26,856	2,381	52,218	6,639	1,798	1,426

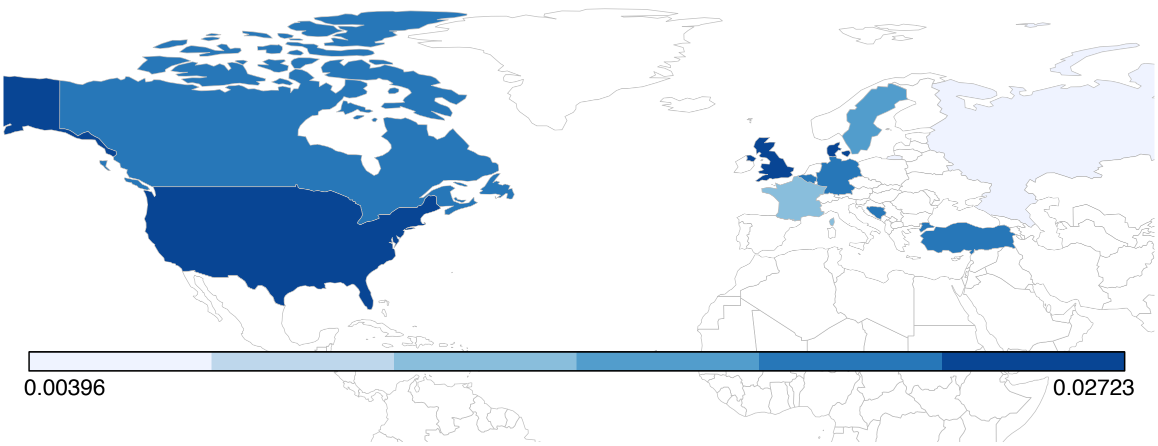
Standard errors in parentheses, clustered at the user level.
The table reports results when calculated using a window length of [-1,+1].
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.13: Terrorist attacks and anti-Muslim and far-right rhetoric, by number of attacks in a country

Days before and after attack	[-1,1]	[-7,7]	[-30,30]
	Anti-Muslim content (sd units)		
After attack = 1	0.074***	0.018**	0.010**
	(0.017)	(0.007)	(0.004)
Number of attacks in a country	-0.086***	0.027***	-0.009***
	(0.002)	(0.001)	(0.001)
After attack = 1 × Number of attacks in a country	-0.002	0.002*	0.001
	(0.002)	(0.001)	(0.001)
Constant	0.617***	-0.207***	0.052***
	(0.016)	(0.007)	(0.004)
User fixed effects	✓	✓	✓
R^2	0.186	0.148	0.144
Number of clusters	11,787	13,547	14,444
Number of observations	230,890	971,219	3,250,107
	Far-right content (sd units)		
After attack = 1	-0.002	-0.015**	-0.009**
	(0.016)	(0.007)	(0.004)
Number of attacks in a country	-0.077***	0.018***	-0.017***
	(0.002)	(0.001)	(0.001)
After attack = 1 × Number of attacks in a country	0.015***	0.009***	0.004***
	(0.002)	(0.001)	(0.001)
Constant	0.570***	-0.127***	0.104***
	(0.016)	(0.007)	(0.004)
User fixed effects	✓	✓	✓
R^2	0.099	0.065	0.053
Number of clusters	11,070	12,963	14,090
Number of observations	226,935	936,334	3,119,581

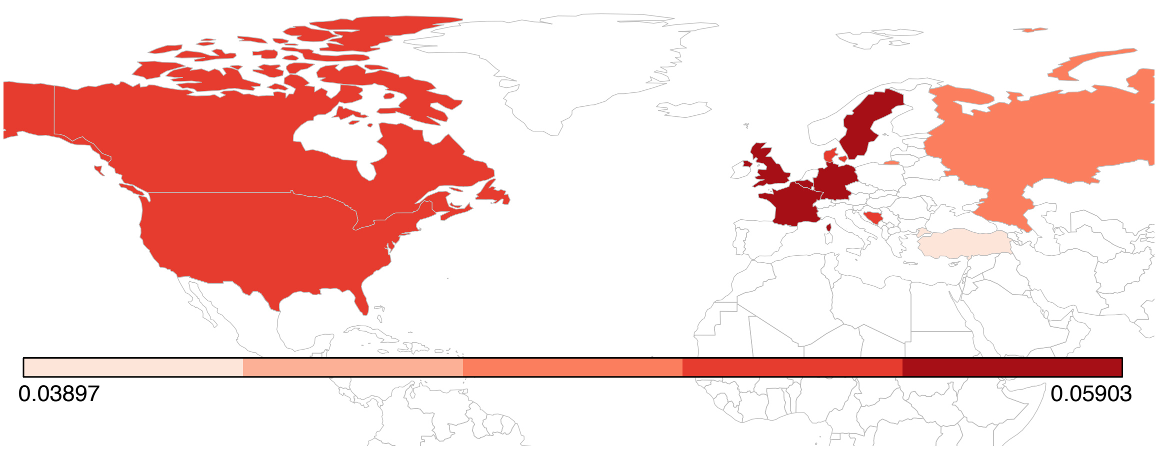
Standard errors in parentheses, clustered at the user level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure 3.7: Anti-Muslim similarity scores by country



Note: The map shows the country-level averages of the anti-Muslim similarity scores for countries targeted by terrorism and included in the analysis.

Figure 3.8: Far-right similarity scores by country



Note: The map shows the country-level averages of the far-right similarity scores for countries targeted by terrorism and included in the analysis.

Chapter 4

Do Community Engagement Efforts Reduce Extremist Rhetoric on Social Media?

Tamar Mitts[1]

[1]This research was approved by the Columbia University Institutional Review Board under protocol IRB-AAAQ7657. I thank Shaked Doron for her research assistance.

Abstract

Over the past few years, efforts at countering violent extremism (CVE) have increased around the world. In the United States, much of the focus has been on community engagement – programs aiming to reduce radicalization by empowering local communities to identify warning signs of extremism before individuals engage in violence. The emphasis on community engagement is rooted in the idea that local knowledge held by families, neighbors, and friends is crucial for countering radicalization. Understanding whether engaging communities is effective is of paramount importance, especially with the rising accessibility of extremist materials on the Internet and social media. However, to date, there has been little systematic study of the effectiveness of community engagement programs in reducing radicalization in the United States. This chapter uses new geo-located data on the online behavior of Islamic State supporters and their followers on Twitter, along with information on community engagement activities held by the Department of Homeland Security's Office for Civil Rights and Civil Liberties during the Obama Administration from 2014 to 2016, to examine whether community engagement events are associated with reductions in pro-ISIS content on Twitter in these localities. The findings show that community engagement activities are followed by a decrease in online pro-ISIS rhetoric, especially in areas that have held a large number of these events.

4.1 Introduction

Countering radicalization and support for violent extremism is becoming a central policy area around the world. The rise of Islamic State (ISIS) and its ability to recruit individuals via online propaganda and social networks has intensified efforts to find solutions to extremist violence (Vidino and Hughes, 2015*a*). Since 2011, tens of thousands have radicalized in support for ISIS, some seeking to become foreign fighters and others attempting to plot terror acts in their home countries (Schmitt and Sengupta, 2015). What can be done to mitigate this wave of violence and extremism? Are there certain policies that governments can implement to counteract the narrative that Islamic State and other groups promote among their followers? While there are several possible responses to extremism,[2] this chapter focuses on one specific strategy: engaging communities in thwarting radicalization in the United States. Community engagement has been a central counter-extremism policy pursued by governments around the world, particularly since 9/11 (Briggs, 2010; Challgren et al., 2016; Romaniuk, 2015).

Engaging communities is rooted in earlier models of community policing developed in the 1990s.[3] Unlike professional policing that focuses on the response and prosecution of crime, community policing emphasizes crime prevention by addressing specific needs of local communities and by involving citizens in police activities (Cordner, 2014). For example, instead of relying solely on the police to actively intervene and put a stop to drug dealing, community members assist by monitoring and reporting the activity of drug dealers in their neighborhood. In a similar manner, engaging communities in countering extremism is based on the idea that extremist violence can be prevented by involving local communities

[2]For example, countering extremist propaganda online (Fernandez, 2015) or initiating educational programs among vulnerable populations (Aldrich, 2014)

[3]In the United States, community policing was formally enacted in the United States in the 1994 Violent Crime Control and Law Enforcement Act. see `https://www.congress.gov/bill/103rd-congress/house-bill/3355`

in efforts to detect early signs of extremism.

This logic relies on several assumptions. First, that radicalization is a process that begins with a cognitive stage in which an individual embraces an extremist ideology, which culminates in behavioral manifestations of radicalization like committing violence (Neumann, 2013; Sedgwick, 2010). Some scholars do not agree with this assumption, arguing that not all extremist behaviors are preceded by the adoption of radical ideologies (Borum, 2011). Many, however, agree that radicalization follows a sort of continuum from mere ideology to actual violence (Horgan, 2008; Wilner and Dubouloz, 2010; Crossett and Spitaletta, 2010).

Second, models of community engagement rest on the premise that the cognitive phase of radicalization can be detected by people who are close to a radicalizing individual. As members of local communities tend to have close personal connections, they are most likely to notice changes in behavior and thus serve as "early warning systems" of violent extremism (Briggs, 2010). Third, proponents of community engagement assume that individuals who experience cognitive radicalization can be swayed away from the radicalization path by members of their community. By addressing local communities' concerns and engaging in activities that reinforce a sense of belonging in the broader society, governments believe that they can counter the "us versus them" narrative that extremist groups promulgate in their recruitment propaganda (Executive Office of the President of the United States, 2011).

While many agree that stopping violent extremism is of paramount importance, not everyone believes that government-sponsored community engagement is the right way to go. Civil rights and Muslim advocacy groups strongly criticize efforts to engage Muslim communities in countering radicalization. The Council on American-Islamic Relations (CAIR), for example, has argued that government-sponsored community engagement is likely to be ineffective, as community figures working with the government are viewed as not credible by radicalizing individuals. In addition, CAIR stressed that community engagement efforts unjustly focus on Muslim communities, even though far-right extremist violence has led to far greater casualties than violence by Islamic extremists (Council on American-Islamic Relations, 2016). The American Civil Liberties Union (ACLU) has stressed that involving Muslim communities in countering extremism pressures community members to monitor each other, which can harm community cohesion and can easily lead to government overreach (American Civil Liberties Union, 2016*a*). Indeed, focusing only on the Muslim community stigmatizes Muslims and reinforces Islamophobic stereotypes, which can be counterproductive from a counter-radicalization standpoint (Patel and German, 2015).

For these reasons, it is of crucial importance to study the effectiveness of community engagement to counter extremism. However, empirically evaluating these efforts is a challenging task. One reason is that community engagement events are not randomized; focusing on the behavior of a very small minority in society, they almost always selectively target specific communities. Moreover, an 'effective' counter-radicalization program requires obtaining some sort of measurable evidence of a "decrease" in radicalization. This is challenging, as observing the absence of radical sympathies does not necessarily imply a decline but may simply reflect the absence of a tendency toward extremist ideologies from the outset. Third, many counter-extremism efforts are not reported to the public,[4] which makes it challenging to systematically study them.

In this chapter, I take advantage of publicly reported community engagement events held by the Department of Homeland Security's Office of Civil Rights and Civil Liberties (CRCL) from 2014 to 2016, and combine them with high-frequency, geo-located panel data on tens of thousands of individuals who follow Islamic State accounts on Twitter, to examine

[4]In fact, in February 2016, ACLU filed a lawsuit under the Freedom of Information Act against the Department of Homeland Security, Department of Justice, Federal Bureau of Investigation, Office of the Director of National Intelligence, Department of State, Department of Health and Human Services, and the Department of Education for not releasing records of their countering violent extremism activities. See: `https://www.aclu.org/sites/default/files/field_document/cve_foia_complaint_2.10.16.pdf`

whether community engagement activities are systematically associated with changes in pro-ISIS rhetoric at the local level. The use of high-frequency Twitter data helps overcome the challenge of 'non-evidence' in evaluating counter-extremism efforts, as it serves as a continuous measure of online rhetoric of individuals at risk of radicalization. Analyzing over a hundred community engagement events in a Difference-in-Differences design, I show that community engagement activities are followed by a significant decrease in online pro-ISIS chatter, especially in localities in which CRCL has held a large number of events. The next section describes in detail the context of the study, and the role of community engagement events in the Obama Administration's strategy to counter violent extremism in the United States.

4.2 Countering violent extremism in the United States

In August 2011, the Obama Administration initiated a counter radicalization strategy, "Empowering Local Partners to Prevent Violent Extremism in the United States," to prevent extremist violence in its territories (Obama, 2011). The plan focused on three main areas. First, it sought to increase and strengthen the government's engagement with local communities whose members may be targeted by violent groups. This effort was based on the notion that community members with personal connections to radicalizing individuals — for example, teachers, friends, or family members — are best positioned to detect changes in behavior that might convey early signs of extremism. Building relationships and trust with local communities was seen as important to accessing crucial information on individuals at early stages of radicalization and to provide an opportunity for communities to give feedback on the government's CVE efforts. When setting out its strategy, the Obama Administration stated:

> *"Engagement is essential for supporting community-based efforts to prevent violent extremism because it allows government and communities to share informa-*

> *tion, concerns, and potential solutions. Our aims in engaging with communities to discuss violent extremism are to (1) share sound, meaningful, and timely information about the threat of radicalization to violence with a wide range of community groups and organizations, particularly those involved in public safety issues; (2) respond to community concerns about government policies and actions; and (3) better understand how we can effectively support community-based solutions."*[5]

In addition to community engagement, the initiatives focused on increasing training for government and law enforcement on preventing radicalization and extremism, and seeking ways to counter the propaganda spread by violent groups on the Internet and social media.

The significance that the American government placed on finding solutions to violent extremism increased with the rise of Islamic State, its vast online propaganda machine, and its efforts to recruit foreign fighters around the world (Vidino and Hughes, 2015*a*). Even though individuals had radicalized in America prior to the rise of ISIS, the pace at which the group attracted supporters was unprecedented compared to prior conflicts. From 2014 to 2016, over a hundred individuals have been charged in the U.S. with criminal behavior related to Islamic State (Vidino and Hughes, 2015*b*). Activities that led to charges included providing material support to ISIS and its affiliates, traveling or planning to travel to Syria to become foreign fighters, or plotting violent attacks in the territories of the United States (Greenberg, 2016). While those who displayed 'behavioral radicalization' in the United States are only a tiny minority, security agencies estimate that the number of people who sympathize with ISIS's ideology — those who display 'cognitive radicalization' — is much larger, possibly in the thousands (Vanden Brook, 2015).

In order to prevent radicalization and violence, the United States CVE strategy seeks to support locally-based activities that can sway individuals who are at the early stage of radicalization from the path of extremism (Bjelopera, 2012; Challgren et al., 2016). In late

[5] Obama (2011), p. 5

2014, the government launched a "Three City Pilot" program in three cities in the United States: Boston, Los Angeles and Minneapolis–St. Paul in order to create local solution to ISIS-inspired radicalization. Recommendations from the program included, among other things, increasing local communities' understanding of extremism with training, enhancing collaboration between communities and law enforcement, and building networks between public and private groups to counter extremism (Challgren et al., 2016; Vidino and Hughes, 2015*a*).

At the federal level, several agencies have been tasked with implementing the government's CVE strategy.[6] In this chapter, I focus on the activities of the Department of Homeland Security. The Office of Civil Rights and Civil Liberties (CRCL) in the Department of Homeland Security is responsible for the Department's community engagement efforts. Its goals include "promoting the respect of civil rights and civil liberties in policy creation and implementation" and "communicating with individuals and communities whose civil rights and civil liberties may be affected by Department activities" (U.S. Department of Homeland Security, 2016). In the past few years, CRCL has been holding various community engagement events across the United States to facilitate relationships with local communities and to enhance counter extremism efforts. These activities include (Office for Civil Rights and Civil Liberties, 2016):

- **Community roundtables.** Events that bring together government officials from the federal, state, and local level and leaders from American Arab, Muslim, South Asian, Middle Eastern, Somali, Sikh, Latino, Jewish, an Asian/Asian Pacific Islander communities, in order to strengthen relationships and engagement.
- **Consultation with communities on CVE.** Events in which CRCL representatives

[6] Specifically, the Federal Government tasked the Department of Homeland Security, the National Counter Terrorism Center, and Federal Bureau of Investigation, and the Department of Justice implement the governments countering violent extremism strategy (Executive Office of the President of the United States, 2011).

meet with local communities to share information, discuss community concerns related to extremism, and receive community input on the effects of the Department's polices on the ground.

- **Community awareness briefing.** Meetings in which community members and law enforcement officials are presented with information on the process of radicalization and foreign fighter recruitment by violent groups, in order to increase awareness and knowledge of the phenomenon.
- **Community resilience exercise.** Events in which law enforcement and local communities participate in a half-day exercise designed to build trust and communication, and to empower communities against violent extremism. The training includes discussion of a hypothetical scenario of violent activity, and evaluates the way in which it affects law enforcement and community members. The exercise concludes with developing a local plan to prevent extremism.

Many community engagement meetings held by CRCL do not exclusively focus on extremism. Instead, they cover a wide range of issues related to the Department's activities. The collaboration between communities and the government is meant to facilitate a "shared sense of belonging" among communities and government officials, which arguably helps undermine extremist propaganda (Executive Office of the President of the United States, 2011).

4.3 Criticism of the U.S. CVE program

That said, many civil rights and Muslim advocacy organizations across the United States, such as the American Civil Liberties Union (ACLU) and the Council on American-Islamic Relations (CAIR), have strongly opposed the American CVE strategy on several grounds. First, they argue that the program is ineffective. Community engagement events sponsored

Figure 4.1: Community roundtables in Atlanta, GA and Phoenix, AZ

Photo credit: Islamic Speakers Bureau, Atlanta, Islamic Community Center of Phoenix.[7]

by the government are not likely to be viewed as credible by individuals attracted to extremist propaganda. As the ideology promoted by ISIS and other groups intentionally calls for fighting the American government, initiatives stemming from the government are likely to be viewed with suspicion (Council on American-Islamic Relations, 2015).

In addition, these groups argue that United State's CVE program lacks clear leadership, receives attention only after terrorist attacks, and tends to depend on the whim of local authorities (Council on American-Islamic Relations, 2016; Patel and German, 2015). Since Muslim communities are already targeted with hate crimes and Islamophobia, especially after terrorist attacks, efforts of CVE programs can further a sense of alienation and hostility that can feed into grievances capitalized upon by groups like Islamic State. Finally, critics of the United State's CVE strategy argue that the program is based on a false model of radicalization, which assumes that violent behavior can be predicted by the expression of certain beliefs. However, numerous empirical studies have shown that there is no single path to radicalization (American Civil Liberties Union, 2016*b*; Patel and German, 2015).

The third argument against the American counter-extremism strategy claims that it is unjust. Even though the majority of casualties since 9/11 have been caused by far-right terrorism, counter-radicalization efforts almost always target Muslims (Council on American-Islamic Relations, 2016; Patel and German, 2015). Critics argue that the exclusive focus on Muslim communities stigmatizes Muslims and Islam, and tends to encourage anti-Muslim sentiment. In addition, community engagement events are not as benign as they might seem, because government agencies use them for spying and intelligence gathering. This sort of infiltration of community spaces solely on the basis of religion is unfair, according to civil rights activists (American Civil Liberties Union, 2016*b*).

Furthermore, community engagement frequently tasks members with monitoring each other's behavior. Such monitoring can create a climate of fear, discourage free expression of political opinions, and can be used by the government to suppress dissent (American Civil Liberties Union, 2016*b*; Patel and German, 2015). Finally, critics of CVE argue that the strategy imposes heavy costs on Muslim communities by harming community cohesion, increasing suspicion among members, and by framing community relations with the government only on the basis of security issues (American Civil Liberties Union, 2016*b*; Patel and German, 2015).

These are all very important considerations that should be taken into account when

[7] `http://archive.constantcontact.com/fs036/1101532851599/archive/1108633466286.html`, `http://iccpaz.com/dhs-hosts-crcl-roundtable-at-the-iccp/`

evaluating counter-radicalization policies. To date, however, there has not been a systematic empirical analysis of the link between community engagement events and observable measures of pro-ISIS radicalization. This study brings together new geo-located data on the online behavior of individuals who follow Islamic State accounts on Twitter, who might be at risk of radicalization,[8] along with data on community engagement events held during the Obama Administration by the Department of Homeland Security's Office of Civil Rights and Civil Liberties. I use these new sources of data to examine whether there is a link between community engagement activities and pro-ISIS rhetoric on Twitter.

This is an important step forward in understanding the potential effects of community engagement, as there has been very little systematic empirical evaluations of these programs, especially in the United States. I should note, however, that while this study sheds light on the possible impact of community engagement on the behavior of ISIS followers online, it does not allow concluding that changes in pro-ISIS chatter reflect a decline in radicalization. It is equally possible that these events reduce pro-ISIS chatter by suppressing political expression. Below, I describe the data collection, research design, and results, and discuss several tests to examine the alternative explanation that community engagement events might be discouraging online expression.

4.4 Data

In this section, I describe the data collection for this study. First, I describe how I collected information on the timing and location of community engagement events held by the Department of Homeland Security's Office of Civil Rights and Civil Liberties in various

[8] Online measures of pro-ISIS rhetoric can plausibly proxy underlying support for extremism. Among more a hundred individuals charged in the United States with providing material support for ISIS or plotting a violent attack on the organization's behalf, about 63% used social media when they were radicalizing, and among those, 86% expressed their support for ISIS in publicly viewable posts. This study uses large amounts of publicly viewable posts by ISIS supporters in the United States to measure support for violent extremism. See Table 2.25 in Chapter 2 for more details.

locations across the United States. Second, I describe how I collected data on the online rhetoric of individuals located in the United States who follow Islamic State accounts on Twitter. Finally, I discuss how I matched ISIS followers to community engagement events taking place in their areas based on granular information on their geographic location.

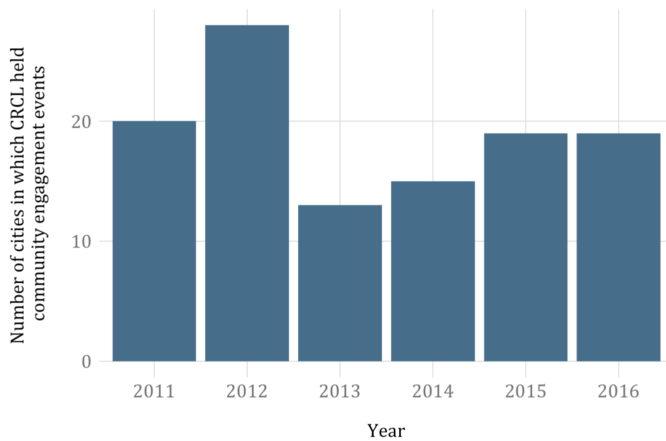
4.4.1 Community engagement events

The Office of Civil Rights and Civil Liberties (CRCL) in the Department of Homeland Security has been holding community engagement events since 2003 (Bjelopera, 2012). The first event took place in Dearborn, Michigan, and community activities soon expanded to other cities in the United States (Schlanger, 2011). In the end of 2010, CRCL began publishing monthly newsletters in which it provided information on its community engagement activities:

> *"This is the first of CRCL's new monthly newsletters. Our goal is to inform members of the public about the Office's activities, including how to make complaints; ongoing and upcoming projects; opportunities to offer comments and feedback ... Public engagement with diverse American communities plays a key role in the DHS mission to protect America while preserving our freedoms ... We are hard at work expanding our engagement program, building a strong stakeholder network of community-based organizations across the country – this newsletter is a part of that effort."* (Schlanger, 2011)

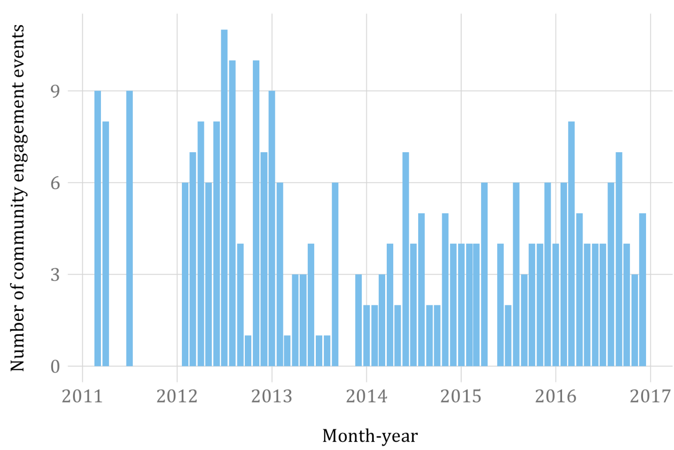
I collected information on all events held by CRCL using these monthly reports, which began providing systematic information on events in 2011. I gathered data on the dates of these events, the cities in which they took place, and the type of engagement activity carried out in each event. These included community roundtables, community awareness briefings, and community resilience exercises, among others. Figure 4.2 displays the number of cities in which CRCL held community engagement activities since 2011. Figure 4.3 shows the number of community engagement events by month. Table 4.9 in the Appendix provides detailed information on each event.[9]

Figure 4.2: Number of cities in which CRCL held community engagement events, by year



Note: The figure presents the number of cities in which The Office for Civil Rights and Civil Liberties (CRCL) in the Department of Homeland Security held community engagement events each year since 2011.

Figure 4.3: Number of CRCL community engagement events, by month



Note: The figure presents the number of community engagement events held each month by The Office for Civil Rights and Civil Liberties (CRCL) in the Department of Homeland Security since 2011.

[9]In this chapter, I focus on events that took place from 2014 to 2016. Thus, Table 4.9 describes these events in detail.

4.4.2 Islamic State supporters on Twitter

To evaluate the possible impact of these community engagement events on the behavior of individuals attracted to ISIS's ideology, I used original Twitter data on Islamic State supporters in the United States, which comes from a larger database on ISIS-affiliated accounts that I collected as part of this dissertation. Below, I provide an overview of the data collection procedure, which included (i) identifying accounts of Islamic State supporters on Twitter and downloading information on their posting history, (ii) coding the extent to which their posts reflected extremist ideology, and (iii) predicting their geographic location using network data. See Chapter 2 for more details on the data collection method.

Identifying Islamic State accounts on Twitter

First, I identified about 15,000 accounts of Islamic State activists — accounts that actively disseminated ISIS propaganda online — that were flagged for suspension from Twitter by the group Controlling Section (@CtrlSec). Controlling Section has been monitoring, since 2015, Twitter accounts identified with ISIS and publicly flagging them for suspension. I downloaded every available piece of information on these accounts before they were suspended from Twitter, including user-level data such as profile picture, description, and self-described location, as well as complete historical tweet timelines. In addition to the core list of ISIS activists, I collected user-level data and tweeting history for all the followers of these accounts, which amount to about 1.6 million users. The followers group includes individuals who follow one or more ISIS activist accounts.

Measuring online expression of extremist ideology

Using the historical tweet timelines for these accounts, I measured the extent to which each tweet represented pro-ISIS content. Specifically, I used supervised machine learning to classify tweets in four different languages (English, Arabic, French, and German) into one or more of the categories listed below.[10]

1. *Travel to Syria or foreign fighters* - tweets describing interest or intent to travel to Syria, and/or discussion of foreign fighters
2. *Sympathy with ISIS* - expressions of support or sympathy with the Islamic State, its ideology and its activities in territories under its control
3. *Life in ISIS territories* - tweets describing the life of ISIS activists in the territories controlled by the Islamic State
4. *Syrian war*- tweets describing events in the Syrian civil war and/or discussion/analysis of those events
5. *Anti-West* - anti-West rhetoric, criticizing Western countries' foreign policy and military operations in the Middle East

I asked human coders from two crowdsourcing platforms, Amazon Mechanical Turk and Crowdflower, to manually label a training set of randomly selected Twitter posts in Arabic, English, French and German. Each tweet was labeled by three coders, and and label(s) were retained for a given tweet only if at least two out of the three coders assigned the same label(s) to the tweet.[11] Using the labeled training set, I predicted the content of all unlabeled tweets using the elastic-net generalized linear model (Friedman, Hastie and Tibshirani, 2010), where the regularization parameter λ was selected by cross-validation. The algorithm employed information on the words in each labeled tweet to 'learn' the categorization rules to classify unlabled tweets. Chapter 2 provides information on the coding procedure and model performance.

[10] I also coded whether tweets represented discourse on Islam in general, but in this chapter I focus on topics that can capture pro-ISIS rhetoric.

[11] The coders were proficient in the languages of the tweets that they labeled.

Predicting ISIS supporters' geographic locations

In order to estimate where ISIS activists and followers are located, I employed a spatial label propagation algorithm developed by Jurgens (2013).[12] This algorithm predicts users' geographic location using geo-location information available in the network, along with information on network structure and the strength of ties between users. Since a very small share of users enabled geo-tagging of their tweets or provided location information in their accounts, I predicted the geographic locations for all users to avoid relying on the small selected subset of users with reported locations. Chapter 2 provides more details on the method, along with information on its prediction accuracy and stability.

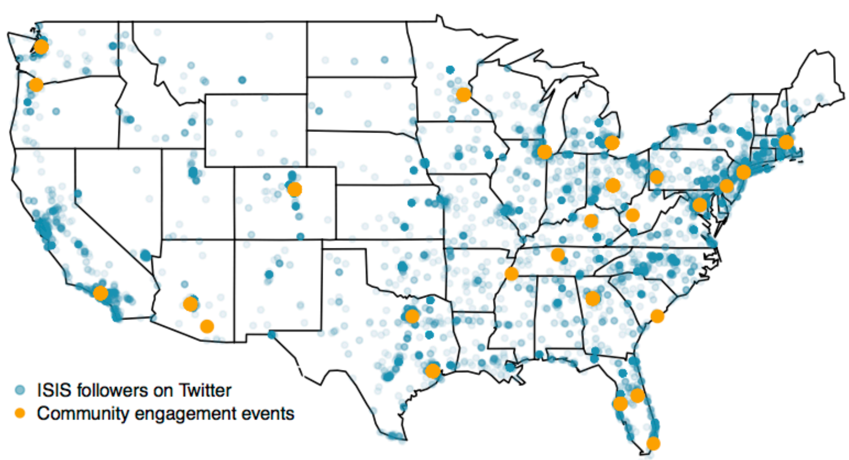
As the goal of this study is to match ISIS supporters to community engagement events held by the Department of Homeland Security, I used the predicted geographic coordinates of users to determine which were located in the United States. Thus, in this chapter, I analyze the online rhetoric of about 47,000 ISIS-affiliated accounts predicted to be located in America, examining whether they changed their pro-ISIS rhetoric in the aftermath of community engagement events. Specifically, and as described in more detail in Section 4.5, I use Difference-in-Differences estimations for each community engagement event to compare changes in pro-ISIS rhetoric by individuals located in the area of the event to changes in such rhetoric by all other individuals. Figure 4.4 displays the predicted locations of these accounts (blue dots), along with the locations of CRCL community engagement events from 2014 to 2016 (orange dots).

Table 4.1 provides summary statistics on the tweeting patterns of ISIS sympathizers in the United States. The top panel shows the number of tweets that each user posted on each topic from 2014 to 2016. For example, ISIS followers in America posted an average of about

[12] The use of location prediction in social network research is a growing field (for example, see Backstrom, Sun and Marlow (2010); McGee, Caverlee and Cheng (2013); Jurgens et al. (2015)). To the best of my knowledge, this study is one of the first applications of these methods in the study of Islamic State online networks.

9 tweets discussing travel to Syria or foreign fighters, with a maximum of 230 tweets, and an average of 5.5 tweets expressing sympathy with ISIS, with a maximum of 145 tweets.[13] The bottom panel provides additional information on these accounts, such as whether they were flagged for suspension by the group Controlling Section (about 0.2%), whether they were already suspended (2.6%), and the number of ISIS-affiliated accounts that each user followed (the mean being about 3 accounts, with a maximum of 1,793).

Figure 4.4: ISIS supporters and CRCL Community Engagement Activities in the United States



Note: The figure plots the predicted locations of accounts of ISIS activists and their followers in the United States (blue dots). In addition, it shows the locations in which CRCL held community engagement events from 2014 to 2016 (orange dots).

4.4.3 Creating datasets for each community engagement event

In order to facilitate analysis of the relationship between community engagement activities and pro-ISIS rhetoric on Twitter, I created separate datasets of Twitter posts produced by ISIS sympathizers in the United States, around each of 112 engagement events held by

[13]In Table 4.1, a tweet was coded as belonging to a topic if its predicted value from the classification model described in section 4.4.2 was above the mean of the predicted values for this category.

Table 4.1: Summary statistics for ISIS supporters in the United States

Statistic	N	Mean	St. Dev.	Min	Max
Travel to Syria or foreign fighters (#)	35,248	8.824	19.092	0	230
ISIS sympathy (#)	35,248	5.597	12.358	0	145
Life in ISIS territories (#)	35,248	8.946	19.719	0	262
Syrian war (#)	35,248	4.153	9.467	0	118
Anti-West (#)	35,248	4.721	10.513	0	151
All topics (#)	35,248	21.666	46.075	0	561
Flagged as an ISIS activist (0/1)	47,296	0.002	0.041	0	1
Suspended by Twitter (0/1)	47,287	0.026	0.159	0	1
ISIS accounts following (#)	47,287	2.970	17.431	0	1,793

Note: The table reports summary statistics for accounts of ISIS activists and followers located in the United States. The number of tweets for each topic reflect tweets that coded 1 if their predicted value of belonging to the topic (i.e., sympathy with ISIS, life in ISIS territories, travel to Syria or foreign fighters, Syrian war, or anti-West) was above the mean of the predicted values for that topic, and 0 if not.

the Department of Homeland Security from 2014 to 2016. For each event, I identified the tweets posted by ISIS sympathizers in the 7, 14, 21, and 30 days before and after the event. I created two binary indicators to (i) differentiate between posts appearing before and after each event, as well as (ii) distinguish between tweets posted by individuals located in or out of the area of the event. Specifically, for each community engagement event, I created the variable *Post*, which is coded 1 when a tweet appeared after the event and 0 otherwise, and a variable *In event area*, which is coded 1 when a tweet was posted by an individual located in the area of the event and 0 otherwise. Finally, to quantify the extent to which each post expressed pro-ISIS rhetoric, I used the predicted values generated for each tweet by the classification model described in section 4.4.2, for each of the five content categories. Table 4.9 in The Appendix provides summary statistics for each of these 112 datasets. Each row represents a different dataset for a different community engagement event, and the columns show the distribution of the variables described above in each of the 112 datasets.

4.5 Research design and results

Since this study analyzes the relationship between pro-ISIS rhetoric and over a hundred community engagement events, it is impractical to present regression results for each event separately. To uncover patterns underlying all events held by the Department of Homeland Security from 2014 to 2016, I employ two types of analysis. First, I conduct a pooled analysis where I examine all community engagement events simultaneously. Second, I carry out meta analysis of the results of individual events, as described in detail below. Meta analysis is a useful tool for the purpose of this study, as it allows systematically evaluating the relationship between community engagement events and online pro-ISIS rhetoric when considering many events simultaneously. In addition, it enables examining how the relationship varies as a function of event characteristics. In the first part of this section, I describe the Difference-in-Differences model I used to analyze community engagement events when all events are pooled together. In the second part, I describe the meta analysis method I employed to evaluate the overall impact of 112 community engagement events on the rhetoric of ISIS sympathizers in the United States.

4.5.1 Identification strategy

The key identifying assumption in this Difference-in-Differences design is that in the absence of a community engagement event, individuals located in the event area and individuals who do not would follow parallel trends in their online expression of pro-ISIS rhetoric. While it is possible that community engagement events target specific areas that might be more prone to have individuals "at risk" of radicalization (Obama, 2011), the *over-time changes* in pro-ISIS online posting should not be significantly different between the groups before the occurrence of community engagement events.

To empirically test this identification assumption, I visually examine whether the two groups display parallel trends before CRCL community engagement events. Figure 4.5 plots pre- and post-time trends in pro-ISIS rhetoric — a standardized variable capturing online posts on all topics: (i) travel to Syria or foreign fighters, (ii) ISIS sympathy, (iii) life in ISIS territories, (iv) Syrian war, and (v) anti-West — for the group of individuals located in event areas (black) and the group of those who do not (gray). The x-axis is the number of days between the date of a community engagement event and the date in which ISIS followers posted on Twitter. In order to observe time trends for all events simultaneously, the figure normalizes, for all CRCL events, the difference in days between community engagement events and the timing of Twitter posts. I calculated the average pro-ISIS content by each of the two groups in each day, and applied nonparametric smoothing piecewise to the pre- and post- time periods for each group, using a Gaussian kernel with a bandwidth of 30 days.

Figure 4.5 shows that the trends in pro-ISIS rhetoric over time for individuals located in event areas and those who do not are parallel prior to the day in which CRCL held community engagement events. Only after community engagement events we observe a shift in those trends, where pro-ISIS content by individuals in event areas decreases, but the rhetoric of those outside of event areas does not change. Interestingly, we also observe that the average pro-ISIS rhetoric is overall higher for individuals located in event areas. This suggests that CRCL may intentionally target locations that might have greater numbers of individuals at risk of radicalization.

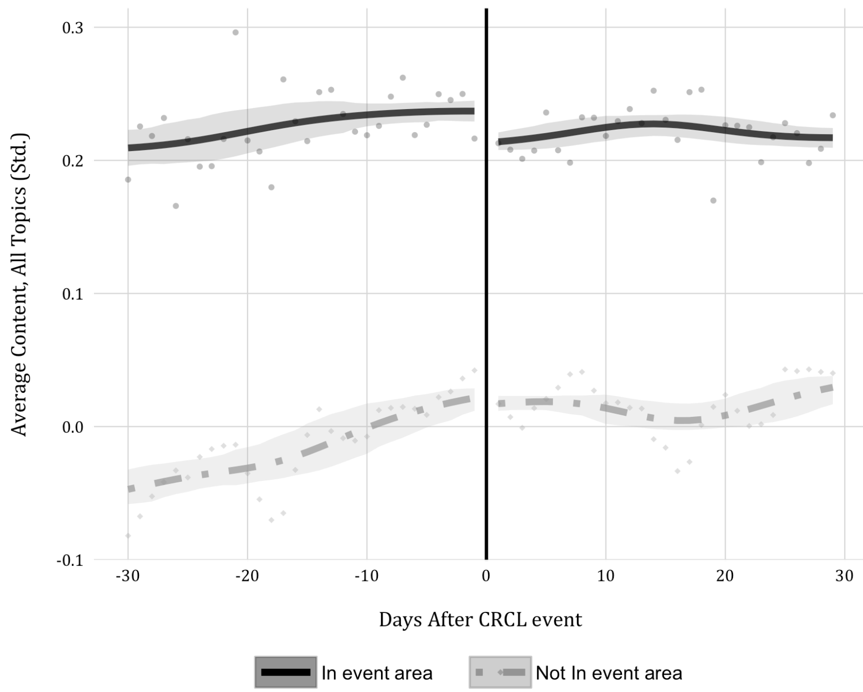
4.5.2 Difference-in-Differences model

To examine the relationship between community engagement events and pro-ISIS rhetoric on Twitter, I estimate the following least squares regression model:

$$y_{i,j,k} = \beta_1 Post_{i,j,k} + \beta_2 In\ event\ area_{i,j,k} + \beta_3 (Post_{i,j,k} \times In\ event\ area_{i,j,k}) + \alpha_k + \varepsilon_j \quad (4.1)$$

where $y_{i,j,k}$ is the predicted value of a given topic (i.e., travel to Syria or foreign fighters, sympathy with ISIS, life in ISIS territories, Syrian war, anti-West) for tweet i posted by

Figure 4.5: Pro-ISIS rhetoric: Parallel trends



Note: The Figure presents pre- and post-time trends in pro-ISIS rhetoric — a standardized variable capturing online posts on all topics: (i) travel to Syria or foreign fighters, (ii) ISIS sympathy, (iii) life in ISIS territories, (iv) Syrian war, and (v) anti-West — for the group of individuals located in event areas (black) and the group of those who do not (gray). The x-axis is the number of days between the community engagement event date and the day in which ISIS followers posted on Twitter. I calculated the average pro-ISIS content by each of the two groups in each day, and applied nonparametric smoothing piecewise to the pre- and post-time periods for each group, using a Gaussian kernel with a bandwidth of 30 days.

user j surrounding event k; *Post* is an indicator coded 0 where a tweet appears before event k and 1 afterwards; *In event area* is an indicator coded 1 when a tweet was posted by an individual who is predicted to be located in the area of event k, and 0 otherwise; and α_k is an event fixed effect. In this specification, β_3 is the Difference-in-Differences coefficient of interest, reflecting how the change in pro-ISIS rhetoric after community engagement events is different for individuals located in event areas, compared to the change in rhetoric of users outside event areas. Standard errors are clustered at the user level to account for serial correlation in tweet content posted by the same user. All outcome variables are standardized.

4.5.3 Pooled analysis results

Table 4.2 reports the results estimated from a pooled Difference-in-Differences analysis of the relationship between 112 community engagement events and pro-ISIS rhetoric on social media, captured 7, 14, 21, and 30 day before and after each event. The results encompass data from 32,694,069 (7-day results), 64,848,887 (14-day results), 100,860,573 (21-day results), and 147,372,103 (30-day results) geo-located tweets generated from 2014 to 2016 by individuals who follow Islamic State accounts on Twitter. The coefficient on the interaction term, *Post* $\times$ *In event area*, represents the Difference-in-Differences coefficient of interest.

Overall, the findings show that community engagement events are systematically associated with reductions in pro-ISIS rhetoric on Twitter. Each of the six panels in Table 4.2 reports the results for a different content category. It can be seen that the Difference-in-Differences coefficient, *Post* $\times$ *In event area*, is negative and statistically significant at the 1% to 5% level for almost all categories. When analyzing data from 30 days before and after the events and considering all topics together (Table 4.2, Panel 1, Column 4), the results show that community engagement events are linked to a reduction of about 4 percentage points in pro-ISIS rhetoric among individuals located in event areas. This is a relative decrease of 18% over the baseline coefficient of 0.233 standard deviations, which is the pre-post decrease in pro-ISIS content among individuals located in areas where CRCL held community engagement events.

The findings also hold when considering the content categories separately. For example, posting on the topic 'Travel to Syria or foreign fighters' decreases by about 3 percentage points after community engagement events in areas where they take place, which a relative decrease of 16% over the baseline coefficient of 0.189 standard deviations. Similarly, discourse on the topic 'ISIS sympathy' decreases by about 2 percentage points (a relative decrease of 26% over the baseline) and posting on 'Life in ISIS territories' is reduced by about 3 percentage points (a relative decrease of 13% over the baseline). Tweets discussing the Syrian civil war decrease by about 2 percentage points after community engagement events in areas where they take place, which is a relative decrease of 22% over the baseline of 0.1 standard deviations. Interestingly, the results for anti-West rhetoric, which includes a lot of anti-America tweets,[14] do not significantly change after community engagement events among individuals located in event areas. When considering shorter time windows, such as 14 and 21 days before and after community engagement events, the results hold as well, but the coefficient estimates have slightly smaller magnitudes.

[14]For examples, see Table 2.2 in Chapter 2.

Table 4.2: Pooled Diff-in-Diff Analysis: Community engagement and pro-ISIS rhetoric on Twitter

	(1) 7 Days	(2) 14 Days	(3) 21 Days	(4) 30 Days
		1. All Topics		
Post	-0.007***	0.008***	0.028***	0.049***
	(0.001)	(0.001)	(0.001)	(0.001)
In event area	0.204***	0.215***	0.228***	0.233***
	(0.011)	(0.009)	(0.009)	(0.008)
Post × In event area	**-0.024***	**-0.039*****	**-0.033*****	**-0.042*****
	(0.014)	(0.011)	(0.008)	(0.007)
R^2	0.01	0.009	0.008	0.006
		2. Travel to Syria or foreign fighters		
Post	-0.005***	0.005***	0.017***	0.033***
	(0.001)	(0.001)	(0.000)	(0.001)
In event area	0.167***	0.171***	0.181***	0.189***
	(0.007)	(0.006)	(0.006)	(0.005)
Post × In event area	**-0.014**	**-0.023*****	**-0.026*****	**-0.032*****
	(0.01)	(0.007)	(0.006)	(0.005)
R^2	0.005	0.004	0.004	0.003
		3. ISIS sympathy		
Post	-0.002***	0.004***	0.012***	0.021***
	(0.001)	(0.000)	(0.000)	(0.000)
In event area	0.052***	0.058***	0.065***	0.067***
	(0.007)	(0.005)	(0.005)	(0.005)
Post × In event area	**-0.011**	**-0.016****	**-0.013****	**-0.018*****
	(0.01)	(0.007)	(0.005)	(0.005)
R^2	0.002	0.002	0.002	0.001
		4. Life in ISIS territories		
Post	-0.006***	0.005***	0.024***	0.043***
	(0.001)	(0.001)	(0.001)	(0.001)
In event area	0.245***	0.255***	0.268***	0.271***
	(0.011)	(0.008)	(0.008)	(0.007)
Post × In event area	**-0.012**	**-0.031*****	**-0.029*****	**-0.034*****
	(0.013)	(0.01)	(0.007)	(0.007)
R^2	0.008	0.007	0.006	0.005
		5. Syrian war		
Post	-0.007***	0.003***	0.016***	0.028***
	(0.001)	(0.001)	(0.000)	(0.001)
In event area	0.08***	0.088***	0.096***	0.1***
	(0.008)	(0.006)	(0.006)	(0.005)
Post × In event area	**-0.016**	**-0.022*****	**-0.017*****	**-0.022*****
	(0.011)	(0.008)	(0.006)	(0.006)
R^2	0.004	0.003	0.003	0.002
		6. Anti-West		
Post	-0.004***	0.003***	0.013***	0.022***
	(0.001)	(0.001)	(0.000)	(0.001)
In event area	0.007	0.02***	0.026***	0.028***
	(0.008)	(0.006)	(0.005)	(0.005)
Post × In event area	**0.023**	**0.001**	**0.000**	**-0.005**
	(0.016)	(0.01)	(0.007)	(0.006)
R^2	0.003	0.002	0.002	0.002

Note: The table reports coefficients estimated from a pooled Difference-in-Differences analysis of the relationship between 112 community engagement events and pro-ISIS rhetoric on social media, captured 7, 14, 21, and 30 day before and after each event. The results encompass data from 32,694,069 (7-day results), 64,848,887 (14-day results), 100,860,573 (21-day results), and 147,372,103 (30-day results) geo-located tweets generated from 2014 to 2016 by individuals who follow Islamic State accounts on Twitter. All outcome variables are standardized. The analysis includes event fixed effects and clustered standard errors (reported in parentheses) at the user level. * p<0.10, ** p<0.05, *** p<0.01.

4.5.4 Meta analysis

Another approach to examine whether community engagement activities influence pro-ISIS rhetoric online is to carry out meta analysis of the Difference-in-Differences estimations of individual events.[15] Meta-analysis is "the statistical synthesis of the data from separate but similar, i.e. comparable studies, leading to a quantitative summary of the pooled results" (Last et al., 2001). I treat the results from each community engagement event as a separate 'study,' in order to quantitatively estimate their combined effect. As the underlying analysis for individual events uses regression models, I follow the recommendations of existing research on meta-analysis of regression slopes, and use standardized ("beta") coefficients when carrying out the meta analysis (Becker and Wu, 2007; Peterson and Brown, 2005).

Model

For each content category (i.e., travel to Syria or foreign fighters, sympathy with ISIS, life in ISIS territories, Syrian war, anti-West), I estimated 112 Difference-in-Differences regressions, corresponding to different community engagement events held by the Department of Homeland Security's Office for Civil Rights and Civil Liberties in various locations in the United States. The assumption behind meta analysis models is that each individual estimate corresponds to a true latent coefficient, measured with some error:

$$y_i = \theta_i + \varepsilon_i \tag{4.2}$$

In the equation above, y_i represents the β_3 coefficient from Equation (1) for study i; θ_i represents the (unknown) true coefficient; and ε_i is a sampling error, assumed to be distributed normally with mean 0 and variance v_i (this is without loss of generality by the

[15]Specifically, I estimate the following Difference-in-Differences model for each event: $y_{i,j} = \beta_1 Post_{i,j} + \beta_2 \textit{In event area}_{i,j} + \beta_3(Post_{i,j} \times \textit{In event area}_{i,j}) + \varepsilon_j$ This is the same model as Equation in (4.1), but it does not include event fixed effects, as each event is estimated separately. In addition, in order to facilitate the meta analysis, all variables in these models (independent and dependent variables) are standardized.

Central Limit Theorem). If N_i is the sample size of the ith study, the sampling variance is calculated as follows:

$$v_i = \frac{\left(1 - y_i^2\right)^2}{N_i - 1} \tag{4.3}$$

In the meta analysis, I assume that the sample of Islamic State supporters in the United States represents the same underlying population, and thus I estimate fixed effects meta analytic models (Patall and Cooper, 2008; Viechtbauer et al., 2010). This assumption is reasonable, since the same population is being estimated over and over again for each event. The only change between events is the categorization of tweets as being 'in the event area' or not, or being posted before or after the event.[16] The fixed-effects model estimates the underlying true average effect using weighted least squares:

$$\bar{\theta}_w = \frac{\sum_{i=1}^{k} w_i \theta_i}{\sum_{i=1}^{k} w_i} \tag{4.4}$$

In the equation above, $\bar{\theta}_w$ represents the weighted average of the true latent coefficients estimated for each community engagement event (θ_i), where the weight is inverse-proportional to the sampling error: $w_i = \frac{1}{v_i}$. In other words, the model gives more weight to studies with smaller sampling variance.

4.5.5 Meta analysis results

Table 4.3 reports the meta analysis results for 112 community engagement events held between 2014 and 2016. The coefficients represent the weighted average of the coefficients estimated for each event (i.e., $\bar{\theta}_w$ from equation (4)), measured in standard deviation units, where the time window surrounding each event is set to 30 days before and 30 days after

[16] This may raise the concern that the same tweet i might be coded differently in different estimations. For example, in the estimation of event A, tweet i might be coded 1 for the variable *In event area*, but in the regression of event B is it coded as 0. At worst, this misclassification will bias the results of event B towards zero (by making the outcomes for the treatment and control groups more similar to each other, on average), but it is important to preserve the same population for the meta analysis.

the event. As in the pooled analysis, we find that community engagement events are systematically associated with reductions in pro-ISIS rhetoric on Twitter.

Since the coefficients reported in Table 4.3 are measured in standard deviation units, which are somewhat hard to interpret substantively, I report in Table 4.4 the percent change reflected in each Difference-in-Differences coefficient. The percent change is calculated as the coefficient β_3 (*Post* $\times$ *In event area*) divided by the coefficient β_2 (*In event area*), and reflects the change in pro-ISIS content after community engagement events for individuals located in event areas, compared to pro-ISIS content generated in these areas before the events. Table 4.4 also reports the results when the impact of the events are measured with different time windows: 7, 14, and 21 days.

As in the pooled analysis, the results in Table 4.4 show that the relationship between community engagement events and pro-ISIS rhetoric on Twitter is overall negative, and becomes more strongly negative as the window around the event expands. For example, when comparing the content of tweets in the 7 days before and after community engagement events, discourse on foreign fighters decreases by about 5%, but the difference is only marginally significant with a p-value of 0.09. The percent change for other topics is also negative, but not statistically significant. For the 14 and 21 day estimations, community engagement events are associated with a significant decrease of 5-8% in discourse on foreign fighters, 10-13% decrease in tweets expressing sympathy with ISIS, 4-5% decrease in tweets describing life in ISIS territories, and 8-10% decrease in discussion of the Syrian war. In these estimations, the anti-West topic has a positive percent change, but it is not statistically significant.

Finally, the results are strongest in terms of magnitude and significance when considering estimations using the 30 day window. The rightmost column in Table 4.4 shows that 30 days after community engagement events discussion on foreign fighters decreased by more than 11%, tweets sympathizing with ISIS decreased by almost 19%; discussion of life in ISIS territories decreased by 8%, and tweets describing the Syrian war decreased by 14%. Unlike the pooled analysis, the meta analysis results for the 30-day window show that anti-West content significantly decreased by almost 17% after community engagement events in areas where they were held.

Taken together, results from the pooled and meta analyses show, systematically across over a hundred community engagement events and tens of thousands of individuals, that engagement activities are followed by a decrease in pro-ISIS rhetoric. At least during the Obama Administration, these events were meant to share information, give feedback, and build trust between communities and the government. The finding that the impact of these events is strongest after 30 days might mean that it takes communities time to identify, counsel, and help people who show signs of radicalization in their areas. If one assumes that reduced pro-ISIS rhetoric on Twitter reflects a decline in radicalization, then these results suggest that engaging communities in countering extremism might be effective. To the best of my knowledge, this is one the first systematic examinations of the possible impacts of community engagement activities in the United States. Much of the CVE strategy implemented in America so far has not been based on rigorous empirical research (Vidino and Hughes, 2015*a*). This study reveals important patterns relating to these initiatives that prior work has not been able to observe.

Table 4.3: Meta Analysis: Community engagement and pro-ISIS rhetoric on Twitter

	Estimate	Std. Err.	P-value
		1. All topics	
Post	2.36***	0.01	0.00
In event area	0.78***	0.01	0.00
Post × In event area	-0.09***	0.01	0.00
Intercept	0.00	0.01	0.71
		2. Travel to Syria or foreign fighters	
Post	1.56***	0.01	0.00
In event area	0.65***	0.01	0.00
Post × In event area	-0.07***	0.01	0.00
Intercept	0.00	0.01	0.81
		3. ISIS sympathy	
Post	1.01***	0.01	0.00
In event area	0.22***	0.01	0.00
Post × In event area	-0.04***	0.01	0.00
Intercept	0.00	0.01	0.86
		4. Life in ISIS territories	
Post	2.08***	0.01	0.00
In event area	0.97***	0.01	0.00
Post × In event area	-0.08***	0.01	0.00
Intercept	0.00	0.01	0.74
		5. Syrian war	
Post	1.31***	0.01	0.00
In event area	0.35***	0.01	0.00
Post × In event area	-0.05***	0.01	0.00
Intercept	0.00	0.01	0.83
		6. Anti-West	
Post	1.07***	0.01	0.00
In event area	0.10***	0.01	0.00
Post × In event area	-0.02**	0.01	0.04
Intercept	0.00	0.01	0.86

Note: The table shows coefficients estimated from a meta analysis of the relationship between 112 community engagement events and pro-ISIS rhetoric on social media, captured 30 day before and after each event. The results encompass data from 147,141,409 geo-located tweets generated from 2014 to 2016 by individuals who follow Islamic State accounts on Twitter. All variables are standardized. Coefficients are multiplied by 100 for presentation purposes.

Table 4.4: Meta Analysis: Community engagement and pro-ISIS rhetoric on Twitter (different time windows)

	7 Days		14 Days		21 Days		30 Days	
	% Change	P-value	% Change	P-value	% Change	P-value	% Change	P-value
Travel to Syria/foreign fighters	-4.88*	0.09	-5.26**	0.01	-7.76***	0.00	-11.33***	0.00
ISIS sympathy	-10.99	0.23	-10.72*	0.08	-12.97***	0.00	-18.51***	0.00
Life in ISIS territories	-2.08	0.26	-3.96***	0.00	-5.35***	0.00	-8.01***	0.00
Syrian war	-8.97	0.18	-9.73**	0.01	-8.48***	0.00	-14.26***	0.00
Anti-West	32.93*	0.07	2.83	0.83	0.77	0.93	-16.97**	0.04
All topics	-4.42*	0.07	-5.91***	0.00	-7.13***	0.00	-11.86***	0.00

Note: The % Change reflects the change in pro-ISIS content after community engagement events for individuals located in event areas, compared to pro-ISIS content in these areas before the events. In technical terms, it represents the meta analysis results for the Diff-in-Diff coefficients ($Post \times In\ event\ area$) across 112 community engagement events, divided by the coefficient *In event area* (where $Post = 0$).

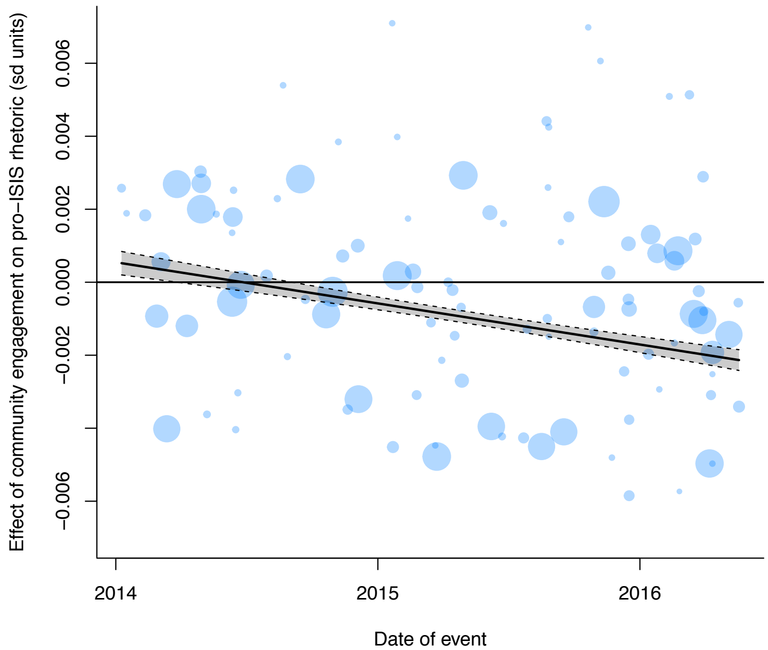
4.5.6 Heterogeneity

In this section, I expand the prior analysis by looking at how the results vary with event-level characteristics, such as the timing of the event, the number of engagement activities held in a given location before the event, as well as the type of event. To examine how the estimated coefficients of community engagement events vary over time, I regressed the estimates of these events on the dates in which they took place. Figure 4.6 presents a meta-analytic scatterplot, in which the observed estimates for community engagement events (measured as a standardized index combining all content categories) are plotted against the date of the event. The resulting regression line has a negative and statistically significant slope, indicating that pro-ISIS rhetoric decreased more strongly after community engagement events taking place in the later part of 2015 and 2016.

Next, I examine how the Difference-in-Differences coefficients vary with the number of events held in each area. If engaging communities in countering extremism is effective, then more events might lead to a greater reduction in pro-ISIS tweets by ISIS sympathizers in these localities. Figure 4.7 shows a meta analytic scatterplot of the estimated coefficients for each event, plotted against the number of community engagement activities taking place in each location prior to the event. Here, too, we find a negative and statistically significant relationship, which might suggest that more community engagement activities held in an area might be more effective for countering online support for extremism.

Finally, in Table 4.5, I summarize the heterogeneous results we find when considering different event types. The table reports, for each type of activity (community roundtable, community resilience exercise, and community awareness briefing), the estimated pooled Difference-in-Differences coefficients in the first row, and in the second row, the estimated difference from the pooled Difference-in-Differences result for each event type. The results show that community roundtables — events in which government representatives and members of various communities meet to strengthen relationships and engagement —

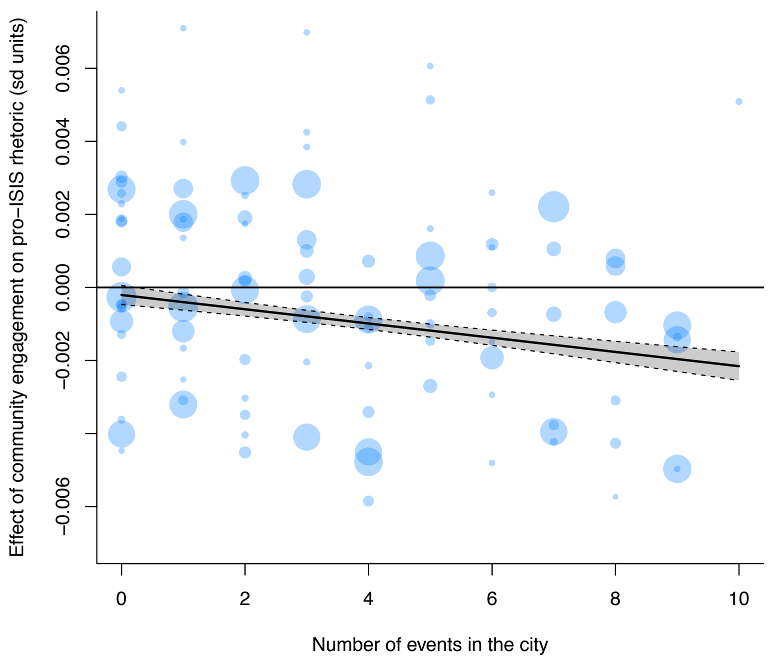
Figure 4.6: The effect of community engagement events on pro-ISIS rhetoric, by date of event



Note: The figure presents the meta-analytic scatterplot of the observed effects estimated for individual community engagement events, where the dependent variable is a standardized index of all topics, calculated 30 days away from the event. The x-axis plots the date in which each event was taking place. The point sizes are proportional to the inverse of the standard errors, which means that events with larger samples have larger points. The predicted average effects are included (with corresponding 95% confidence intervals), calculated from the meta-analysis model.

are associated with a stronger negative coefficient of community engagement on pro-ISIS rhetoric when compared to all other event types. A similar finding is reported for community awareness briefings — events in which community members and law enforcement officials are presented with information on the process of radicalization and foreign fighter recruitment. The results for community resilience exercises also have a negative relationships but the results are not statistically significant.

Figure 4.7: The effect of community engagement events on pro-ISIS rhetoric, by number of events in a city



Note: The figure presents the meta-analytic scatterplot of the observed effects estimated for individual community engagement events, where the dependent variable is a standardized index of all topics, calculated 30 days away from the event. The x-axis plots the number of community engagement events held in each city at the time of each event. The point sizes are proportional to the inverse of the standard errors, which means that events with larger samples have larger points. The predicted average effects are included (with corresponding 95% confidence intervals), calculated from the meta-analysis model.

4.6 Do community engagement events suppress expression?

One primary objection to the results described above is that the findings do not reflect a decline in radicalization, but the suppression of expression. As described in section 4.3, community engagement activities might discourage individuals from expressing their opinion, views, and beliefs, as they facilitate a climate of fear by tasking community members with monitoring each other's behavior (American Civil Liberties Union, 2016*b*). Thus, the observed decline in pro-ISIS rhetoric after community engagement activities might be driven by individuals' abstention from expressing their opinions on Twitter.

Table 4.5: Community engagement and pro-ISIS rhetoric, by event type

	Estimate	Std. Err.	P-value
Estimated $\bar{\theta}$	-0.07***	0.02	0.00
Community roundtable	-0.03*	0.02	0.07
Estimated $\bar{\theta}$	-0.09***	0.01	0.00
Community awareness briefing (CAB)	-0.10*	0.06	0.08
Estimated $\bar{\theta}$	-0.09***	0.01	0.00
Community resilience exercise (CREX)	-0.02	0.14	0.89

Note: The table shows coefficients estimated from a meta analysis of the relationship between 112 community engagement events and pro-ISIS rhetoric on social media, captured 30 day before and after each event. The table reports, for each type of activity (community roundtable, community resilience exercise, and community awareness briefing), the estimated pooled Difference-in-Differences coefficients in the first row, and in the second row, the estimated difference from the pooled Difference-in-Differences result for each event type. All variables are standardized. Coefficients are multiplied by 100 for presentation purposes.

Relatedly, the lower number of pro-ISIS tweets might be caused by the migration of ISIS sympathizers to private social media platforms. Feeling more strongly monitored by community members, individuals who are interested in Islamic State's ideology might choose to abandon the public Twitter platform altogether. Finally, community engagement might increase surveillance and government intervention in the lives of ISIS sympathizers, which could result in a reduction in their public expression of pro-ISIS sentiment. This might happen, for example, when government agencies request Twitter to suspend accounts of individuals who are accused of supporting extremism. These behaviors will result in an overall reduction in pro-ISIS content, but this decline would not necessarily reflect deradicalization.

To examine these possibilities, I carry out several additional estimations. First, I evaluate whether community engagement events suppress overall expression on Twitter. If they do, we should observe ISIS supporters located in event areas reduce the number of tweets — regardless of their content — after community engagement events. I counted the number of Twitter posts that ISIS sympathizers posted in each locality in the 7, 14, 21, and 30 days before and after each event. I aggregated the tweet-level dataset to a locality-time level dataset, where the locality is a geographic unit (city, town, etc.), and the time is a binary indicator coded 1 for tweet-sums appearing after an event and 0 otherwise.[17] I then estimate the following least squares regression, for each window length:

$$y_{ikt} = \beta_1 \text{Post}_{ikt} + \beta_2 \text{In event area}_{ikt} + \beta_3(\text{Post}_{ikt} \times \text{In event area}_{ikt}) + \alpha_i + \varepsilon_k \qquad (4.5)$$

In the equation above, y_{ikt} represents the number of tweets in location k at time t posted before and after event i; Post_{ikt} is an indicator coded 0 where a tweet-sum is calculated for tweets appearing before event i and 1 afterwards; In event area_{ikt} is an indicator coded 1 for sums of tweets posted by individuals predicted to be located in the area of event i, and 0 otherwise; and α_i is an event fixed effect. As before, β_3 is the Difference-in-Differences coefficient of interest, reflecting how the change in the number of tweets after the event is different for locations in which community engagement events took place, compared to the change the number of tweets in locations where events did not take place. Standard errors are clustered at the locality level.

Table 4.6 shows the results. It can be seen that community engagement events are not systematically associated with changes in the number of tweets posted by ISIS sympathizers in areas where they take place. The interaction term *Post* $\times$ *In event area* is null in all window sizes, and its sign is not consistent. These results suggest that community engagement events are not affecting the number of tweets posted by ISIS supporters on Twitter.

Next, I evaluate whether community engagement events might increase surveillance of ISIS followers on Twitter. While this is difficult to measure, it is possible to use data on

[17] Locality geographical data was taken from the United States' census Topologically Integrated Geographic Encoding and Referencing (TIGER) database. Shape files for geographical units were taken from "Cartographic Boundary Shapefiles - Metropolitan and Micropolitan Statistical Areas and Related Statistical Areas." (`https://www.census.gov/geo/maps-data/data/cbf/cbf_msa.html`)

Table 4.6: Community engagement events and the number of tweets

	7 days	14 days	21 days	30 days
Post	-14.49***	-9.56***	1.66	104.08***
	(3.36)	(2.44)	(1.40)	(23.11)
In event area	254.26***	503.64**	715.11***	900.49***
	(97.51)	(198.11)	(266.85)	(347.68)
Post × In event area	-1.97	27.69	127.95	194.44
	(53.54)	(120.02)	(124.84)	(154.67)
Constant	134.78***	229.19***	326.47***	936.82***
	(28.78)	(49.37)	(71.99)	(225.13)
Event fixed effects	✓	✓	✓	✓
R^2	0.016	0.015	0.015	0.027
Observations	29,970	36,026	39,696	18,488

Note: Standard errors in parentheses, clustered at the locality level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

the suspension rate of ISIS supporters to proxy for increased surveillance. Using user-level data,[18] I created a week-by-week panel data for ISIS followers (i.e., accounts that follow core ISIS accounts), in which I measured (i) whether they were suspended from Twitter and (ii) whether they were flagged for suspension by Controlling Section (@ctrlsec). Replicating the meta analysis in equations (4.2) through (4.4), I estimate the relationship between community engagement events and these outcomes. Due to data availability limitations, I am only able to study thirty community engagement events taking place between January and June of 2016.

Table 4.7 presents the meta analysis results. The coefficients represent the weighted average of the coefficients estimated for each event, measured in standard deviation units, where the time window surrounding each event is set to 30 days before and 30 days after the event. As before, each event was first estimated separately in a Difference-in-Differences regression, where the variable *In event area* differentiated between individuals who were located in the area of a community engagement event and those who did not, and the

[18]These data come from continuous observations user-level data in the ISIS activists/followers database collected by the author, which are refreshed over time as users change their profile data.

variable *Post* differentiated between user-level data observed before and after the event. The coefficient on the interaction term, *Post* $\times$ *In event area*, represents the Difference-in-Differences coefficient combining all events.

The results show that community engagement events are not associated with a greater suspension rate of individuals located in the event area. In panel A in Table 4.7, the coefficient on *Post* $\times$ *In event area* is positive, but not statistically significant. Interestingly, the result is different for the flagging outcome described in Panel B in Table 4.7. Here, we can see that community engagement events are associated with lower flagging rate of accounts of individuals located in event areas. As the flagging of accounts for suspension by Controlling Section (@ctrlsec) is strongly linked to the content that these accounts disseminate (see Table 2.28 in Chapter 2), this suggests that the lower flagging rate is driven by the lower number of pro-ISIS tweets posted in event areas after community engagement events. This findings hold across all window lengths, as can be seen in Table 4.8.

Taken together, the results of this analysis do not support the argument that community engagement events suppress ISIS sympathizers' overall expression on Twitter. Nonetheless, while ISIS supporters located in event areas did not reduce the number of tweets that they posted, they seemed to have have changed their *content*: after community engagement activities, ISIS sympathizers expressed less pro-ISIS rhetoric. This result might be interpreted as a sign of de-radicalization, but it can also be driven by ISIS supporters limiting their expression of specific (e.g., pro-ISIS) topics on Twitter. Similarly, the finding that community engagement events are not systematically associated with greater suspension or flagging of ISIS accounts might be the result of users' greater awareness to monitoring. Thus, while the findings show a systematic decrease in pro-ISIS rhetoric after community engagement events, they do not allow determining the reason behind this decline.

Table 4.7: Meta analysis: Community engagement and account-level changes

	Estimate	Std. Err.	P-value
A. Suspended from Twitter			
Post	2.12***	0.04	0.00
In event area	0.25***	0.04	0.00
Post × In event area	0.03	0.04	0.50
Intercept	0.00	0.04	1.00
B. Flagged as an ISIS activist			
Post	12.68***	0.05	0.00
In event area	0.64***	0.05	0.00
Post × In event area	-0.16***	0.05	0.00
Intercept	-1.23***	0.05	0.00

Note: The table shows coefficients estimated from a meta analysis of the relationship between 30 community engagement events and accounts suspension and flagging, captured 30 day before and after each event. All variables are standardized. Coefficients are multiplied by 100 for presentation purposes.

Table 4.8: Meta analysis: Community engagement and account-level changes (different time windows)

	Change (%)	P-value	Change (%)	P-value	Change (%)	P-value	Change (%)	P-value
Suspended by Twitter	8.20	0.70	6.34	0.70	10.55	0.50	10.55	0.50
Flagged as an ISIS activist	-33.67***	0.00	-35.12***	0.00	-24.66***	0.00	-24.66***	0.00

Note: The % Change reflects the change in suspension and flagging rates after community engagement events for individuals located in event areas, compared to suspension and flagging rates in these areas before the events. In technical terms, it represents the meta analysis results for the Diff-in-Diff coefficients ($Post \times In\ event\ area$) across 112 community engagement events, divided by the coefficient *In event area* (where $Post = 0$).

4.7 Conclusion

Over the past few years, efforts to counter radicalization and violent extremism have increased across the world. In the United States during the Obama Administration, a large portion of CVE activities focused on building trust and engagement with local communities. These initiatives were premised on the idea that community members are best positioned to help radicalizing individuals because of their local-level, context-specific knowledge and expertise. While many initiatives to engage communities have taken place in recent years, there has been little systematic empirical research on how they might affect online extremist behaviors by individuals at risk of radicalization. In this study, I sought to shed light on the impact of community engagement activities held by the Department of Homeland Security's Office of Civil Rights and Civil Liberties (CRCL), by combining granular data on community engagement events with information on the online behavior of Islamic State sympathizers in the United States.

Results from over 100 community engagement activities show that these events were systematically and significantly associated with a reduction in pro-ISIS rhetoric on Twitter among individuals located in event areas. Specifically, the data show that CRCL events were followed by a decrease in discourse on foreign fighters or travel to Syria, reduction in tweets expressing sympathy with ISIS, and a decrease in the number of tweets discussing the Syrian civil war and life in ISIS-controlled territories. The patterns in this study were robust to a large number of community engagement activities and tens of thousands of individuals. However, the results were inconclusive with respect to whether the reduction in pro-ISIS rhetoric was caused by de-radicalization or by the suppression of political expression in these areas. Further research is needed to shed light on this important question.

Overall, this study makes several contributions to existing research. First, by providing granular, geo-located high-frequency data on the online behavior of Islamic State sympathizers in the United States, the study measures an over-time "decrease" in pro-ISIS rhetoric, which could serve as a proxy for radicalization. Online measures of pro-ISIS rhetoric can plausibly reflect underlying support for extremism: a large majority of individuals who radicalized in support for ISIS in the United States have publicly expressed their favorable views towards the organization on social media platforms.[19] As most research on countering extremism has struggled with identifying measures of de-radicalization, this is an important step forward.

Second, this study contributes to current research on community engagement, which is based on sporadic empirical examples, by conducting a rigorous analysis of over 100 community engagement events. The use of multiple events allows generalizing the conclusions beyond specific examples, and enables a more nuanced analysis of the heterogeneity of the findings for different event characteristics. For example, the study found that community engagement events are more effective in areas that hold a larger number of activities, and that specific event types, such community roundtables, are associated with greater reductions in pro-ISIS rhetoric, compared to other types of engagement activities.

Third, the chapter provides a model for future work seeking to study links between local events and online behavior on social media. By predicting the geographic locations of thousands of Islamic State supporters on Twitter, this study was able to incorporate an important geographic dimension to the analysis of social media data that is usually not systematically taken into account. The ability to analyze geo-located high-frequency Twitter data and match it to local activities provides an opportunity to study political behavior in new and exciting ways.

Finally, by focusing on community engagement, this study did not address possible solutions to the crucially important pattern discussed in the first two chapters of this dissertation: that anti-Muslim hostility might be driving radicalization and support for Islamic

[19]See Table 2.25 in Chapter 2 for more information.

State. U.S. government programs have been focused on countering radicalization per se rather than targeting Islamophobia and anti-Muslim sentiment, which, as this dissertation has suggested, may be important drivers of the former. Interventions of this sort might be a fruitful avenue of future academic and policy work.

4.8 Appendix

Table 4.9 provides summary statistics for each of 112 community engagement events taking place in the United States from 2014 to 2016. Each row represents summary statistics for a different dataset, collected for each community engagement event. The table provides information on the timing and location of each event, as well as a summary of the variables in each dataset and the total number of observations. In the table, the minimum and the maximum values of all variables are 0 and 1, respectively. Thus, I report them once for each dataset.

Figure 4.8 presents a cumulative forest plot for each event analyzed in the meta analysis. A cumulative forest plot presents the pooled estimated coefficient as each event's estimate is added to the analysis. The figure shows a forest plot of fixed-effects meta-analysis results for the summary index of pro-ISIS rhetoric, calculated in standard deviation units. Each row represents an estimate for one event. The figure plots 95% confidence intervals for the meta-analysis model, derived from the studies' sampling variances.

Figure 4.9 plots influence diagnostics in the meta analysis. It allows detecting influential cases or outlying studies. Evaluating the results with and without influential cases allows testing the robustness of the meta analysis results. The figure shows that there are several influential events, but event 71 in particular has a strong influence on the results.[20] To examine whether the results are robust to the exclusion of this event, I re-estimated the meta analysis models without event 71. The results are reported in Table 4.10, which shows that the results still hold.

[20]The event was a community roundtable taking place in Phoenix, Arizona on March 30, 2016.

Table 4.9: Summary statistics for each event dataset

			Post		*In event area*		*Travel/FF*		*ISIS sympathy*		*ISIS life*		*Syrian war*		*Anti-West*				
Date	City	State	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	min	max	n
2014-01-09	Atlanta	GA	0.80	0.40	0.0069	0.08	0.21	0.26	0.13	0.28	0.19	0.33	0.10	0.24	0.12	0.26	0.00	1.00	44541
2014-01-16	Los Angeles	CA	0.68	0.47	0.0083	0.09	0.21	0.26	0.13	0.29	0.19	0.34	0.10	0.24	0.12	0.26	0.00	1.00	52174
2014-02-11	Minneapolis	MN	0.52	0.50	0.0002	0.01	0.21	0.26	0.14	0.29	0.19	0.34	0.10	0.24	0.12	0.26	0.00	1.00	73204
2014-02-27	Houston	TX	0.57	0.50	0.0003	0.02	0.22	0.26	0.14	0.29	0.20	0.34	0.10	0.24	0.13	0.27	0.00	1.00	81272
2014-03-05	Phoenix	AZ	0.57	0.49	0.0010	0.03	0.22	0.26	0.14	0.29	0.20	0.34	0.10	0.24	0.13	0.27	0.00	1.00	83487
2014-03-13	Denver	CO	0.56	0.50	0.0057	0.08	0.22	0.26	0.14	0.30	0.20	0.34	0.10	0.24	0.12	0.27	0.00	1.00	85518
2014-03-27	Chicago	IL	0.51	0.50	0.0047	0.07	0.22	0.26	0.14	0.30	0.20	0.34	0.10	0.24	0.12	0.27	0.00	1.00	89707
2014-04-10	Denver	CO	0.47	0.50	0.0063	0.08	0.22	0.26	0.14	0.30	0.20	0.34	0.10	0.24	0.12	0.27	0.00	1.00	90901
2014-04-29	New York	NY	0.49	0.50	0.0143	0.12	0.21	0.26	0.14	0.30	0.19	0.34	0.10	0.24	0.12	0.26	0.00	1.00	85649
2014-04-30	Los Angeles	CA	0.49	0.50	0.0100	0.10	0.21	0.26	0.14	0.30	0.19	0.34	0.10	0.25	0.12	0.26	0.00	1.00	85231
2014-04-30	New York	NY	0.49	0.50	0.0141	0.12	0.21	0.26	0.14	0.30	0.19	0.34	0.10	0.25	0.12	0.26	0.00	1.00	85231
2014-05-08	Tampa	FL	0.49	0.50	0.0004	0.02	0.21	0.26	0.14	0.30	0.19	0.34	0.10	0.24	0.12	0.26	0.00	1.00	85196
2014-05-21	Minneapolis	MN	0.52	0.50	0.0003	0.02	0.22	0.26	0.14	0.30	0.19	0.33	0.10	0.24	0.12	0.27	0.00	1.00	87344
2014-06-12	Atlanta	GA	0.56	0.50	0.0042	0.06	0.22	0.27	0.14	0.30	0.19	0.33	0.10	0.24	0.13	0.27	0.00	1.00	95279
2014-06-12	Houston	TX	0.56	0.50	0.0005	0.02	0.22	0.27	0.14	0.30	0.19	0.33	0.10	0.24	0.13	0.27	0.00	1.00	95279
2014-06-13	Chicago	IL	0.56	0.50	0.0058	0.08	0.22	0.27	0.14	0.30	0.19	0.33	0.10	0.24	0.13	0.27	0.00	1.00	96124
2014-06-14	Houston	TX	0.56	0.50	0.0004	0.02	0.22	0.27	0.14	0.30	0.19	0.33	0.10	0.24	0.13	0.27	0.00	1.00	96951
2014-06-17	New York	NY	0.55	0.50	0.0090	0.09	0.22	0.27	0.14	0.30	0.19	0.33	0.10	0.24	0.13	0.27	0.00	1.00	98862
2014-06-20	Los Angeles	CA	0.55	0.50	0.0112	0.11	0.22	0.27	0.15	0.30	0.19	0.33	0.10	0.24	0.13	0.27	0.00	1.00	101399
2014-06-24	Chicago	IL	0.56	0.50	0.0058	0.08	0.22	0.27	0.14	0.30	0.19	0.33	0.10	0.24	0.13	0.27	0.00	1.00	104752
2014-07-01	Washington	DC	0.57	0.50	0.0098	0.10	0.22	0.27	0.15	0.30	0.19	0.33	0.10	0.24	0.13	0.26	0.00	1.00	109620
2014-07-08	Washington	DC	0.56	0.50	0.0097	0.10	0.22	0.27	0.14	0.30	0.19	0.33	0.10	0.24	0.13	0.27	0.00	1.00	114146
2014-07-29	Washington	DC	0.50	0.50	0.0092	0.10	0.22	0.27	0.14	0.30	0.19	0.33	0.10	0.24	0.12	0.26	0.00	1.00	121074
2014-07-30	Denver	CO	0.49	0.50	0.0022	0.05	0.22	0.27	0.14	0.30	0.19	0.33	0.10	0.24	0.12	0.26	0.00	1.00	121380
2014-08-14	Seattle	WA	0.48	0.50	0.0012	0.03	0.22	0.27	0.14	0.29	0.18	0.33	0.10	0.24	0.12	0.26	0.00	1.00	120742
2014-08-22	Orlando	FL	0.49	0.50	0.0002	0.01	0.22	0.26	0.14	0.29	0.18	0.33	0.10	0.24	0.12	0.26	0.00	1.00	118667
2014-08-27	Los Angeles	CA	0.49	0.50	0.0036	0.06	0.22	0.26	0.14	0.29	0.18	0.33	0.10	0.24	0.12	0.26	0.00	1.00	118656
2014-08-28	New York	NY	0.49	0.50	0.0146	0.12	0.22	0.27	0.14	0.29	0.18	0.33	0.10	0.24	0.12	0.26	0.00	1.00	118582
2014-09-15	Chicago	IL	0.52	0.50	0.0039	0.06	0.22	0.26	0.13	0.29	0.18	0.33	0.10	0.24	0.12	0.26	0.00	1.00	119795
2014-09-22	Columbus	OH	0.52	0.50	0.0011	0.03	0.22	0.26	0.13	0.29	0.18	0.33	0.10	0.24	0.12	0.26	0.00	1.00	122006
2014-10-21	Chicago	IL	0.50	0.50	0.0057	0.08	0.22	0.26	0.14	0.29	0.19	0.33	0.09	0.23	0.12	0.26	0.00	1.00	127545
2014-10-30	Boston	MA	0.50	0.50	0.0041	0.06	0.22	0.26	0.14	0.29	0.19	0.33	0.09	0.24	0.12	0.26	0.00	1.00	129944
2014-11-04	Houston	TX	0.51	0.50	0.0008	0.03	0.22	0.26	0.14	0.29	0.19	0.33	0.10	0.24	0.12	0.26	0.00	1.00	131587
2014-11-07	Minneapolis	MN	0.51	0.50	0.0002	0.01	0.22	0.26	0.14	0.29	0.19	0.33	0.10	0.24	0.12	0.26	0.00	1.00	132080
2014-11-13	Los Angeles	CA	0.51	0.50	0.0052	0.07	0.22	0.26	0.14	0.30	0.19	0.33	0.10	0.24	0.12	0.26	0.00	1.00	132680
2014-11-13	Detroit	MI	0.51	0.50	0.0003	0.02	0.22	0.26	0.14	0.30	0.19	0.33	0.10	0.24	0.12	0.26	0.00	1.00	132680
2014-11-20	Atlanta	GA	0.52	0.50	0.0039	0.06	0.22	0.26	0.14	0.30	0.19	0.33	0.10	0.24	0.12	0.26	0.00	1.00	133515
2014-12-04	Atlanta	GA	0.52	0.50	0.0038	0.06	0.22	0.26	0.14	0.29	0.19	0.33	0.10	0.24	0.12	0.26	0.00	1.00	138936
2014-12-05	Boston	MA	0.52	0.50	0.0036	0.06	0.22	0.26	0.14	0.29	0.19	0.33	0.10	0.24	0.12	0.26	0.00	1.00	139861
2014-12-15	Houston	TX	0.58	0.49	0.0005	0.02	0.22	0.26	0.14	0.30	0.18	0.33	0.09	0.23	0.12	0.26	0.00	1.00	163012
2014-12-18	Tampa	FL	0.59	0.49	0.0018	0.04	0.21	0.26	0.14	0.29	0.18	0.33	0.09	0.23	0.12	0.26	0.00	1.00	170442

Summary statistics for each event dataset (cont.)

			Post		*In event area*		*Travel/FF*		*ISIS sympathy*		*ISIS life*		*Syrian war*		*Anti-West*				
Date	City	State	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	min	max	n
2015-01-21	Seattle	WA	0.56	0.50	0.0027	0.05	0.20	0.26	0.13	0.29	0.17	0.32	0.08	0.22	0.11	0.25	0.00	1.00	246652
2015-01-22	Boston	MA	0.56	0.50	0.0015	0.04	0.20	0.26	0.13	0.29	0.17	0.32	0.08	0.22	0.11	0.25	0.00	1.00	248592
2015-01-28	Chicago	IL	0.53	0.50	0.0034	0.06	0.20	0.26	0.13	0.29	0.16	0.32	0.08	0.21	0.11	0.24	0.00	1.00	261979
2015-01-28	Detroit	MI	0.53	0.50	0.0004	0.02	0.20	0.26	0.13	0.29	0.16	0.32	0.08	0.21	0.11	0.24	0.00	1.00	261979
2015-02-12	Tampa	FL	0.52	0.50	0.0005	0.02	0.20	0.25	0.13	0.28	0.16	0.31	0.08	0.21	0.11	0.24	0.00	1.00	279944
2015-02-19	Denver	CO	0.52	0.50	0.0030	0.05	0.20	0.25	0.12	0.28	0.16	0.31	0.08	0.21	0.11	0.24	0.00	1.00	284135
2015-02-24	Columbus	OH	0.52	0.50	0.0021	0.05	0.20	0.25	0.12	0.28	0.16	0.31	0.08	0.21	0.11	0.24	0.00	1.00	285400
2015-02-25	Phoenix	AZ	0.52	0.50	0.0015	0.04	0.20	0.25	0.12	0.28	0.16	0.31	0.08	0.21	0.11	0.24	0.00	1.00	286965
2015-03-16	New York	NY	0.52	0.50	0.0128	0.11	0.20	0.25	0.12	0.28	0.16	0.31	0.08	0.21	0.10	0.24	0.00	1.00	304927
2015-03-22	Charleston	WV	0.53	0.50	0.0000	0.00	0.20	0.25	0.12	0.28	0.16	0.31	0.08	0.21	0.10	0.24	0.00	1.00	310215
2015-03-24	Minneapolis	MN	0.53	0.50	0.0002	0.01	0.20	0.25	0.12	0.28	0.16	0.31	0.08	0.21	0.10	0.24	0.00	1.00	313268
2015-03-31	Atlanta	GA	0.54	0.50	0.0022	0.05	0.20	0.25	0.12	0.28	0.16	0.31	0.08	0.21	0.10	0.24	0.00	1.00	321090
2015-04-09	Chicago	IL	0.54	0.50	0.0045	0.07	0.20	0.25	0.12	0.28	0.16	0.31	0.08	0.21	0.10	0.24	0.00	1.00	334945
2015-04-15	Houston	TX	0.55	0.50	0.0009	0.03	0.20	0.25	0.12	0.28	0.15	0.31	0.08	0.20	0.10	0.24	0.00	1.00	355272
2015-04-18	Minneapolis	MN	0.55	0.50	0.0001	0.01	0.20	0.25	0.12	0.28	0.15	0.31	0.08	0.20	0.10	0.24	0.00	1.00	361621
2015-04-27	Houston	TX	0.55	0.50	0.0009	0.03	0.19	0.25	0.12	0.28	0.15	0.30	0.08	0.20	0.10	0.23	0.00	1.00	377691
2015-04-28	Los Angeles	CA	0.55	0.50	0.0060	0.08	0.19	0.25	0.12	0.28	0.15	0.30	0.08	0.20	0.10	0.23	0.00	1.00	379824
2015-04-30	Seattle	WA	0.55	0.50	0.0031	0.06	0.19	0.25	0.12	0.28	0.15	0.30	0.08	0.20	0.10	0.23	0.00	1.00	382679
2015-06-06	Phoenix	AZ	0.52	0.50	0.0033	0.06	0.19	0.25	0.12	0.28	0.15	0.30	0.08	0.20	0.10	0.24	0.00	1.00	440159
2015-06-08	Houston	TX	0.52	0.50	0.0007	0.03	0.19	0.25	0.12	0.28	0.15	0.30	0.08	0.20	0.10	0.23	0.00	1.00	441670
2015-06-23	Chicago	IL	0.55	0.50	0.0029	0.05	0.20	0.25	0.12	0.28	0.15	0.31	0.08	0.20	0.10	0.24	0.00	1.00	461330
2015-06-25	Atlanta	GA	0.54	0.50	0.0010	0.03	0.20	0.25	0.12	0.28	0.15	0.31	0.08	0.20	0.10	0.24	0.00	1.00	465762
2015-07-23	Chicago	IL	0.51	0.50	0.0032	0.06	0.19	0.25	0.12	0.28	0.15	0.30	0.08	0.20	0.10	0.23	0.00	1.00	514630
2015-07-28	Lexington	KY	0.51	0.50	0.0000	0.00	0.19	0.25	0.12	0.27	0.15	0.30	0.08	0.20	0.10	0.23	0.00	1.00	520824
2015-08-17	Denver	CO	0.54	0.50	0.0019	0.04	0.19	0.25	0.12	0.27	0.15	0.30	0.08	0.20	0.10	0.23	0.00	1.00	550150
2015-08-24	Philadelphia	PA	0.54	0.50	0.0068	0.08	0.19	0.25	0.12	0.27	0.15	0.30	0.08	0.20	0.10	0.23	0.00	1.00	568133
2015-08-25	Denver	CO	0.54	0.50	0.0018	0.04	0.19	0.25	0.12	0.27	0.15	0.30	0.08	0.20	0.10	0.23	0.00	1.00	572291
2015-08-26	Atlanta	GA	0.54	0.50	0.0010	0.03	0.19	0.25	0.12	0.27	0.15	0.30	0.08	0.20	0.10	0.23	0.00	1.00	575400
2015-08-27	Boston	MA	0.54	0.50	0.0025	0.05	0.19	0.25	0.12	0.27	0.15	0.30	0.07	0.20	0.10	0.23	0.00	1.00	576936
2015-08-27	Los Angeles	CA	0.54	0.50	0.0052	0.07	0.19	0.25	0.12	0.27	0.15	0.30	0.07	0.20	0.10	0.23	0.00	1.00	576936
2015-09-13	Minneapolis	MN	0.53	0.50	0.0002	0.01	0.19	0.25	0.11	0.27	0.14	0.30	0.07	0.20	0.10	0.23	0.00	1.00	629003
2015-09-17	Tampa	FL	0.53	0.50	0.0003	0.02	0.19	0.25	0.11	0.27	0.14	0.30	0.07	0.20	0.10	0.23	0.00	1.00	642335
2015-09-24	Memphis	TN	0.54	0.50	0.0002	0.01	0.19	0.25	0.11	0.27	0.14	0.30	0.07	0.19	0.10	0.23	0.00	1.00	661621
2015-10-08	Boston	MA	0.55	0.50	0.0021	0.05	0.19	0.25	0.11	0.27	0.14	0.30	0.07	0.19	0.10	0.23	0.00	1.00	720641
2015-10-21	Phoenix	AZ	0.57	0.49	0.0011	0.03	0.19	0.25	0.12	0.27	0.14	0.30	0.07	0.19	0.10	0.23	0.00	1.00	814707
2015-10-29	Chicago	IL	0.58	0.49	0.0026	0.05	0.19	0.25	0.12	0.27	0.14	0.30	0.07	0.20	0.10	0.23	0.00	1.00	879395
2015-10-29	Houston	TX	0.58	0.49	0.0002	0.02	0.19	0.25	0.12	0.27	0.14	0.30	0.07	0.20	0.10	0.23	0.00	1.00	879395
2015-11-07	Boston	MA	0.59	0.49	0.0018	0.04	0.19	0.25	0.12	0.27	0.14	0.30	0.07	0.19	0.10	0.23	0.00	1.00	952945
2015-11-12	Minneapolis	MN	0.59	0.49	0.0005	0.02	0.19	0.25	0.12	0.27	0.14	0.30	0.07	0.19	0.10	0.23	0.00	1.00	1002750
2015-11-18	Columbus	OH	0.58	0.49	0.0007	0.03	0.19	0.25	0.12	0.27	0.14	0.30	0.07	0.19	0.10	0.23	0.00	1.00	1086791
2015-11-23	Denver	CO	0.61	0.49	0.0011	0.03	0.19	0.25	0.11	0.27	0.13	0.29	0.07	0.19	0.10	0.22	0.00	1.00	1221192

Summary statistics for each event dataset (cont.)

			Post		*In event area*		*Travel/FF*		*ISIS sympathy*		*ISIS life*		*Syrian war*		*Anti-West*				
Date	City	State	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	min	max	n
2015-12-10	Nashville	TN	0.68	0.47	0.0003	0.02	0.18	0.24	0.11	0.27	0.13	0.29	0.07	0.18	0.09	0.22	0.00	1.00	1800773
2015-12-16	Atlanta	GA	0.68	0.47	0.0008	0.03	0.18	0.24	0.11	0.27	0.13	0.28	0.07	0.18	0.09	0.22	0.00	1.00	1924451
2015-12-16	Miami	FL	0.68	0.47	0.0007	0.03	0.18	0.24	0.11	0.27	0.13	0.28	0.07	0.18	0.09	0.22	0.00	1.00	1924451
2015-12-17	Denver	CO	0.68	0.47	0.0010	0.03	0.18	0.24	0.11	0.27	0.12	0.28	0.07	0.17	0.09	0.22	0.00	1.00	1946166
2015-12-17	Los Angeles	CA	0.68	0.47	0.0026	0.05	0.18	0.24	0.11	0.27	0.12	0.28	0.07	0.17	0.09	0.22	0.00	1.00	1946166
2015-12-17	Tampa	FL	0.68	0.47	0.0005	0.02	0.18	0.24	0.11	0.27	0.12	0.28	0.07	0.17	0.09	0.22	0.00	1.00	1946166
2016-01-13	Detroit	MI	0.50	0.50	0.0003	0.02	0.18	0.24	0.11	0.26	0.12	0.28	0.06	0.17	0.09	0.21	0.00	1.00	2591121
2016-01-16	Columbus	OH	0.51	0.50	0.0004	0.02	0.18	0.24	0.11	0.26	0.12	0.28	0.06	0.17	0.09	0.21	0.00	1.00	2668235
2016-01-25	Denver	CO	0.52	0.50	0.0010	0.03	0.18	0.24	0.11	0.27	0.12	0.28	0.06	0.17	0.09	0.21	0.00	1.00	2856046
2016-01-28	Boston	MA	0.54	0.50	0.0012	0.03	0.18	0.24	0.11	0.27	0.12	0.28	0.06	0.17	0.09	0.21	0.00	1.00	2915099
2016-02-11	Chicago	IL	0.70	0.46	0.0020	0.04	0.17	0.24	0.11	0.26	0.12	0.28	0.06	0.17	0.09	0.21	0.00	1.00	4369777
2016-02-18	Charleston	SC	0.70	0.46	0.0000	0.01	0.17	0.24	0.11	0.26	0.12	0.28	0.06	0.17	0.09	0.21	0.00	1.00	4692288
2016-02-18	Minneapolis	MN	0.70	0.46	0.0002	0.01	0.17	0.24	0.11	0.26	0.12	0.28	0.06	0.17	0.09	0.21	0.00	1.00	4692288
2016-02-23	New York	NY	0.68	0.47	0.0076	0.09	0.17	0.24	0.11	0.26	0.12	0.28	0.06	0.17	0.09	0.21	0.00	1.00	4608500
2016-02-25	Atlanta	GA	0.67	0.47	0.0005	0.02	0.17	0.24	0.11	0.26	0.12	0.28	0.06	0.17	0.09	0.21	0.00	1.00	4577110
2016-02-25	Boston	MA	0.67	0.47	0.0016	0.04	0.17	0.24	0.11	0.26	0.12	0.28	0.06	0.17	0.09	0.21	0.00	1.00	4577110
2016-03-10	Tampa	FL	0.44	0.50	0.0003	0.02	0.18	0.24	0.11	0.27	0.13	0.28	0.07	0.18	0.09	0.21	0.00	1.00	4397808
2016-03-16	Seattle	WA	0.21	0.41	0.0011	0.03	0.18	0.24	0.11	0.27	0.13	0.29	0.07	0.18	0.09	0.22	0.00	1.00	4301821
2016-03-18	New York	NY	0.22	0.41	0.0091	0.10	0.18	0.24	0.11	0.27	0.13	0.29	0.07	0.18	0.09	0.22	0.00	1.00	4265043
2016-03-23	Detroit	MI	0.24	0.43	0.0008	0.03	0.18	0.24	0.11	0.27	0.13	0.29	0.07	0.18	0.09	0.22	0.00	1.00	4189790
2016-03-28	Houston	TX	0.28	0.45	0.0001	0.01	0.18	0.24	0.12	0.27	0.14	0.29	0.07	0.19	0.09	0.22	0.00	1.00	4102509
2016-03-29	Dallas	TX	0.29	0.45	0.0021	0.05	0.19	0.24	0.12	0.27	0.14	0.30	0.07	0.19	0.09	0.22	0.00	1.00	4091249
2016-03-30	Columbus	OH	0.30	0.46	0.0006	0.03	0.19	0.24	0.12	0.27	0.14	0.30	0.07	0.19	0.10	0.22	0.00	1.00	4083441
2016-03-30	Phoenix	AZ	0.30	0.46	0.0010	0.03	0.19	0.24	0.12	0.27	0.14	0.30	0.07	0.19	0.10	0.22	0.00	1.00	4083441
2016-04-07	Atlanta	GA	0.50	0.50	0.0010	0.03	0.19	0.24	0.12	0.28	0.15	0.31	0.08	0.20	0.10	0.23	0.00	1.00	4483780
2016-04-09	Los Angeles	CA	0.58	0.49	0.0050	0.07	0.19	0.25	0.12	0.28	0.15	0.31	0.08	0.20	0.10	0.23	0.00	1.00	4579647
2016-04-11	Denver	CO	0.71	0.46	0.0021	0.05	0.20	0.25	0.13	0.28	0.16	0.32	0.08	0.21	0.10	0.24	0.00	1.00	4364343
2016-04-11	Orlando	FL	0.71	0.46	0.0007	0.03	0.20	0.25	0.13	0.28	0.16	0.32	0.08	0.21	0.10	0.24	0.00	1.00	4364343
2016-04-11	Tampa	FL	0.71	0.46	0.0006	0.03	0.20	0.25	0.13	0.28	0.16	0.32	0.08	0.21	0.10	0.24	0.00	1.00	4364343
2016-05-04	Los Angeles	CA	0.72	0.45	0.0048	0.07	0.21	0.25	0.13	0.29	0.17	0.32	0.08	0.22	0.11	0.24	0.00	1.00	5605973
2016-05-17	Portland	OR	0.15	0.35	0.0006	0.02	0.20	0.25	0.13	0.29	0.17	0.32	0.08	0.22	0.11	0.24	0.00	1.00	5153951
2016-05-18	Seattle	WA	0.10	0.29	0.0022	0.05	0.20	0.25	0.13	0.29	0.17	0.32	0.08	0.22	0.11	0.24	0.00	1.00	5117649

Figure 4.8: Cumulative forest plot

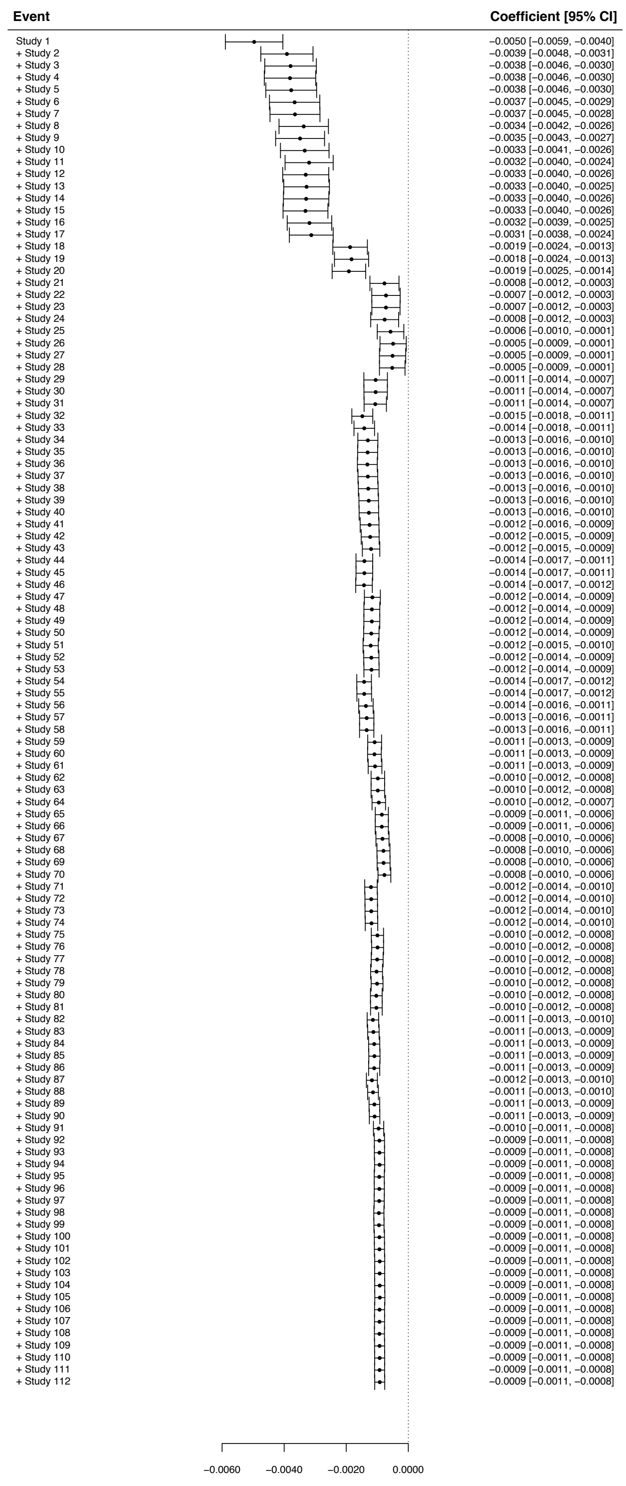


Figure 4.9: Influence of individual events on meta analysis results

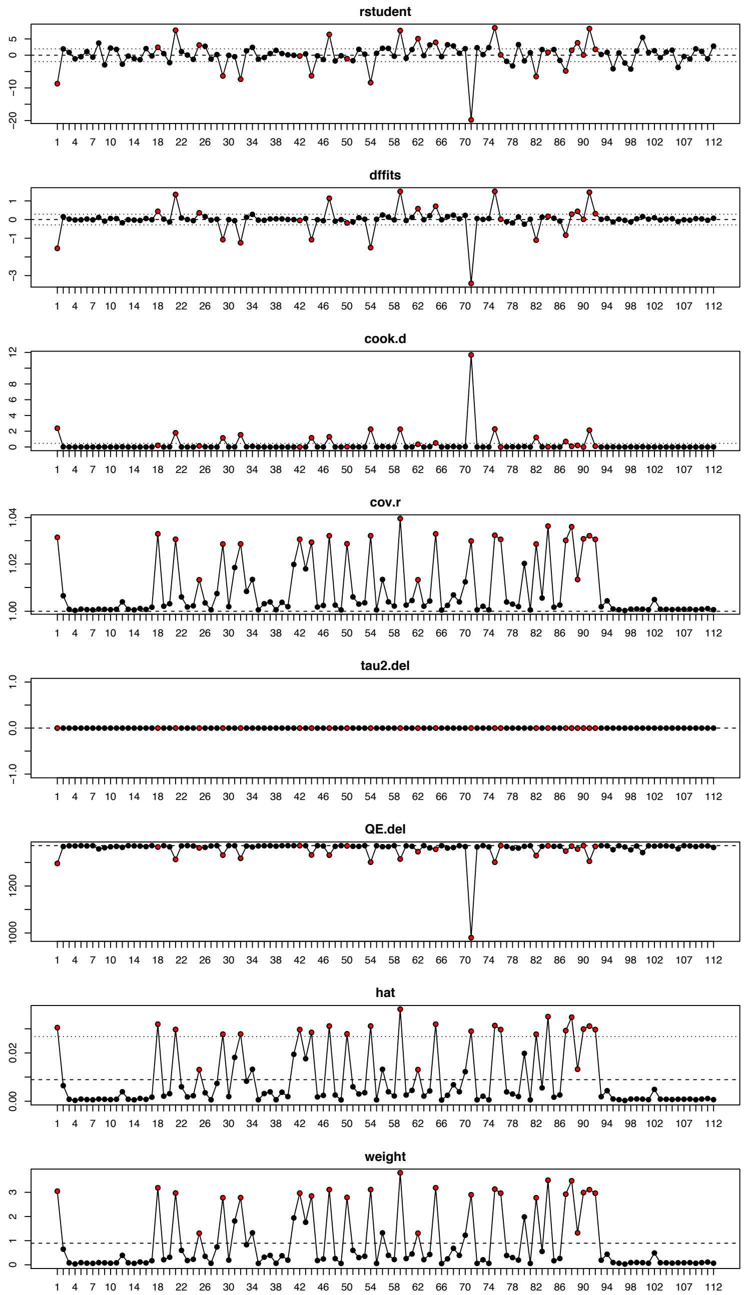


Table 4.10: Meta Analysis: Excluding an influential event

	Estimate	Std. Err.	P-value
	A. Travel to Syria or foreign fighters		
Post	1.38***	0.01	0.00
In event area	0.60***	0.01	0.00
Post × In event area	-0.06***	0.01	0.00
Intercept	0.00	0.01	0.88
	B. ISIS sympathy		
Post	0.90***	0.01	0.00
In event area	0.20***	0.01	0.00
Post × In event area	-0.03***	0.01	0.00
Intercept	0.00	0.01	0.91
	C. Life in ISIS territories		
Post	1.84***	0.01	0.00
In event area	0.90***	0.01	0.00
Post × In event area	-0.05***	0.01	0.00
Intercept	0.00	0.01	0.82
	D. Syrian war		
Post	1.13***	0.01	0.00
In event area	0.32***	0.01	0.00
Post × In event area	-0.03***	0.01	0.00
Intercept	0.00	0.01	0.90
	E. Anti-West		
Post	0.93***	0.01	0.00
In event area	0.08***	0.01	0.00
Post × In event area	-0.00	0.01	0.92
Intercept	0.00	0.01	0.91
	F. All topics		
Post	2.10***	0.01	0.00
In event area	0.71***	0.01	0.00
Post × In event area	-0.06***	0.01	0.00
Intercept	0.00	0.01	0.81

Note: The table shows coefficients estimated from a meta analysis of the relationship between community engagement events and pro-ISIS rhetoric on social media, captured 30 day before and after each event. The results exclude event 71 (a community roundtable in Phoenix, Arizona in March 30, 2016), which was found to be influential in the meta analysis. All variables are standardized. Coefficients are multiplied by 100 for presentation purposes.

Bibliography

Adida, Claire L, David D Laitin and Marie-Anne Valfort. 2014. "Muslims in France: identifying a discriminatory equilibrium." *Journal of Population Economics* 27(4):1039–1086.

Adida, Claire L., David D. Laitin and Marie-Anne Valfort. 2016. *Why Muslim Integration Fails in Christian-Heritage Societies*. Harvard University Press.

Aldrich, Daniel P. 2014. "First Steps Towards Hearts and Minds? USAID's Countering Violent Extremism Policies in Africa." *Terrorism and Political Violence* 26(3):523–546.

American Civil Liberties Union. 2016*a*. "The Problem with "Countering Violent Extremism" Programs.".
URL: *https://www.aclu.org/other/problem-countering-violent-extremism-programs?redirect=problem-countering-violent-extremism-programs*

American Civil Liberties Union. 2016*b*. "What Is Wrong With the Government's "Countering Violent Extremism" Programs.".
URL: *https://www.aclu.org/sites/default/files/field_document/cve_briefing_paper_feb_2016.pdf*

Åslund, Olof and Dan-Olof Rooth. 2005. "Shifts in attitudes and labor market discrimination: Swedish experiences after 9-11." *Journal of Population Economics* 18(4):603–629.

Backstrom, Lars, Eric Sun and Cameron Marlow. 2010. Find me if you can: improving geographical prediction with social and spatial proximity. In *Proceedings of the 19th international conference on World wide web*. ACM pp. 61–70.

Barrett, Richard, Jack Berger, Lila Ghosh, Daniel Schoenfeld, Mohamed el Shawesh, Patrick M Skinner, Susan Sim and Ali Soufan. 2015. "Foreign fighters: An updated assessment of the flow of foreign fighters into Syria and Iraq." *The Soufan Group* 4.

Becker, Betsy Jane and Meng-Jia Wu. 2007. "The synthesis of regression slopes in meta-analysis." *Statistical Science* pp. 414–429.

Becker, Gary S, Yona Rubinstein et al. 2004. "Fear and the response to terrorism: an economic analysis." *University of Chicago mimeo* .

Berger, J and Jonathan Morgan. 2015. "The ISIS Twitter Census: Defining and describing the population of ISIS supporters on Twitter." *The Brookings Project on US Relations with the Islamic World* 3:20.

Berger, JM. 2015. "The Metronome of Apocalyptic Time: Social Media as Carrier Wave for Millenarian Contagion." *Perspectives on Terrorism* 9(4).

Berrebi, Claude and Esteban F Klor. 2008. "Are Voters Sensitive to Terrorism? Direct Evidence from the Israeli Electorate." *American Political Science Review* pp. 279–301.

Biggs, Michael and Steven Knauss. 2012. "Explaining membership in the British National Party: A multilevel analysis of contact and threat." *European Sociological Review* 28(5):633–646.

Bjelopera, Jerome P. 2012. Countering violent extremism in the United States. Congressional Research Service, Library of Congress.

Borum, Randy. 2011. "Radicalization into violent extremism I: A review of social science theories." *Journal of Strategic Security* 4(4):7.

Briggs, Rachel. 2010. "Community engagement for counterterrorism: lessons from the United Kingdom." *International Affairs* pp. 971–981.

Bueno de Mesquita, Ethan and Eric S Dickson. 2007. "The propaganda of the deed: Terrorism, counterterrorism, and mobilization." *American Journal of Political Science* 51(2):364–381.

Burrows, Thomas. 2016. "Islamophobic hate crimes in London have nearly doubled in the last two years as Muslims say they are 'very conscious of keeping your head down'." The Daily Mail.
URL: *http://www.dailymail.co.uk/news/article-3358535/Islamophobic-hate-crimes-London-nearly-doubled-two-years-Muslims-say-conscious-keeping-head-down.html*

Canetti-Nisim, Daphna, Gal Ariely and Eran Halperin. 2008. "Life, pocketbook, or culture: The role of perceived security threats in promoting exclusionist political attitudes toward minorities in Israel." *Political Research Quarterly* 61(1):90–103.

Carter, Joseph A, Shiraz Maher and Peter R Neumann. 2014. *# Greenbirds: Measuring Importance and Influence in Syrian Foreign Fighter Networks*. The International Centre For the Study of Radicalisation and Political Violence.

Challgren, Jonathan, Ted Kenyon, Lauren Kervick, Sally Scudder, Micah Walters, Kate Whitehead, Professors Jeffrey Connor and Carol Rollie Flynn. 2016. "Countering Violent Extremism: Applying the Public Health Model.".

Chatfield, Akemi Takeoka, Christopher G Reddick and Uuf Brajawidagda. 2015. Tweeting propaganda, radicalization and recruitment: Islamic state supporters multi-sided twitter networks. In *Proceedings of the 16th Annual International Conference on Digital Government Research*. ACM pp. 239–249.

Cordner, Gary. 2014. "Community policing." *The Oxford handbook of police and policing* pp. 148–171.

Coryn, Chris L, James M Beale and Krista M Myers. 2004. "Response to September 11: Anxiety, patriotism, and prejudice in the aftermath of terror." *Current Research in Social Psychology* 9(12):165–183.

Council on American-Islamic Relations. 2015. "Brief on countering violent extremism with recommendations for U.S. Government.".

Council on American-Islamic Relations. 2016. "Brief on Countering Violent Extremism (CVE).".
URL: *https://www.cair.com/government-affairs/13063-brief-on-countering-violent-extremism-cve.html*

Crenshaw, Martha. 1981. "The causes of terrorism." *Comparative politics* 13(4):379–399.

Crossett, Chuck and J Spitaletta. 2010. "Radicalization: Relevant psychological and sociological concepts." *The John Hopkins University* .

Dancygier, Rafaela M and David D Laitin. 2014. "Immigration into Europe: Economic discrimination, violence, and public policy." *Annual Review of Political Science* 17:43–64.

Das, Enny, Brad J Bushman, Marieke D Bezemer, Peter Kerkhof and Ivar E Vermeulen. 2009. "How terrorism news reports increase prejudice against outgroups: A terror management account." *Journal of Experimental Social Psychology* 45(3):453–459.

Davila, Alberto and Marie T Mora. 2005. "Changes in the earnings of Arab men in the US between 2000 and 2002." *Journal of Population Economics* 18(4):587–601.

De Figueiredo, Rui JP and Barry R Weingast. 2000. "Vicious cycles: Endogenous political extremism and political violence.".

Dearden, Lizzie. 2015. "Paris attacks: Support rising for far-right Front National as immigration backlash continues." *The Independent*.
URL: *http://www.independent.co.uk/news/world/europe/paris-attacks-support-rising-for-far-right-front-national-as-immigration-backlash-continues-a6745186.html*

ESS Round 7. 2014. "European Social Survey Round 7 Data (2014). Data file edition 2.1." NSD - Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.

Executive Office of the President of the United States. 2011. *Strategic Implementation Plan for Empowering local partners to prevent violent extremism in the United States.* Office of the President.

Fekete, Liz. 2004. "Anti-Muslim racism and the European security state." *Race & Class* 46(1):3–29.

Fernandez, Alberto M. 2015. "Here to stay and growing: Combating ISIS propaganda networks." *The Brookings Project on U.S. Relations with the Islamic World U.S.-Islamic World Forum Papers 2015* .

Fetzer, Joel S and J Christopher Soper. 2005. *Muslims and the state in Britain, France, and Germany.* Cambridge University Press.

Fischer, Peter, Tobias Greitemeyer, Andreas Kastenmüller, Dieter Frey and Silvia Oßwald. 2007. "Terror salience and punishment: Does terror salience induce threat to social order?" *Journal of Experimental Social Psychology* 43(6):964–971.

Flaig, Joseph. 2016. "Khola Hasan from Theydon Garnon says she has experienced increased Islamophobia in Epping Forest." The Guardian.
URL: *http://www.guardian-series.co.uk/news/14459448.Anti_Muslim_media_causing_Islamophobic_atta*

Fortna, Virginia Page. 2015. "Do Terrorists Win? Rebels' Use of Terrorism and Civil War Outcomes." *International Organization* 69(03):519–556.

Friedman, Jerome, Trevor Hastie and Robert Tibshirani. 2010. "Regularization Paths for Generalized Linear Models via Coordinate Descent." *Journal of Statistical Software* 33(1):1–22.
URL: *http://www.jstatsoft.org/v33/i01/*

Front National. 2016. "The National Front Platform.".
URL: *http://www.frontnational.com/le-projet-de-marine-le-pen/*

Getmansky, Anna and Thomas Zeitzoff. 2014. "Terrorism and Voting: The Effect of Exposure to Rockets on Voting in Israeli Elections." *American Political Science Review* (forthcoming).

Goldenberg, Jacob and Moshe Levy. 2009. "Distance is not dead: Social interaction and geographical distance in the internet era." *arXiv preprint arXiv:0906.3202* .

Golder, Matthew. 2016. "Far Right Parties in Europe." *Annual Review of Political Science* 19(1).

Gould, Eric D and Esteban F Klor. 2014. "The Long-Run Effect of 9/11: Terrorism, Backlash, and the Assimilation of Muslim Immigrants in the West." *The Economic Journal* .

Grant, Madeline. 2014. "16% of French Citizens Support ISIS, Poll Finds." Newsweek.
URL: *http://www.newsweek.com/16-french-citizens-support-isis-poll-finds-266795*

Grant, Will J, Brenda Moon and Janie Busby Grant. 2010. "Digital dialogue? Australian politicians' use of the social network tool Twitter." *Australian Journal of Political Science* 45(4):579–604.

Greenberg, Karen J. 2016. "Case by Case: ISIS Prosecutions in the United States." *Center on National Security, Fordham University School of Law* .

Grimmer, Justin and Brandon M Stewart. 2013. "Text as data: The promise and pitfalls of automatic content analysis methods for political texts." *Political Analysis* 21(3):267–297.

Gurr, Ted R. 1970. *Why men rebel*. Princeton University Press.

Hainmueller, Jens and Daniel J Hopkins. 2014. "Public attitudes toward immigration." *Annual Review of Political Science* 17.

Hall, Ellie. 2015. "Gone Girl: An Interview With An American In ISIS." *Buzzfeed News*.
URL: *https://www.buzzfeed.com/ellievhall/gone-girl-an-interview-with-an-american-in-isis?utm_term=.grBaBZdyN#.iv22gLK9N*

Hanes, Emma and Stephen Machin. 2014. "Hate crime in the wake of terror attacks: Evidence from 7/7 and 9/11." *Journal of Contemporary Criminal Justice* .

Healy, Blathnaid. 2015. "Far-right parties gain popularity in Europe after Paris attacks." Mashable.
URL: *http://mashable.com/2015/01/14/far-right-europe-paris-attacks/#w2HcIBkWV8qH*

Homeland Security Committee. 2015. "Final Report of the Task Force on Combating Terrorist and Foreign Fighter Travel.".

Horgan, John. 2008. "From profiles to pathways and roots to routes: Perspectives from psychology on radicalization into terrorism." *The ANNALS of the American Academy of Political and Social Science* 618(1):80–94.

Horgan, John G. 2015. "De-radicalization programs offer hope in countering terrorism.".
URL: *http://www.latimes.com/opinion/op-ed/la-oe-0215-horgan-terrorist-deradicalization-20150215-story.html*

Horowitz, Donald L. 1985. *Ethnic groups in conflict*. University of California Press.

Huddy, Leonie, Stanley Feldman, Charles Taber and Gallya Lahav. 2005. "Threat, anxiety, and support of antiterrorism policies." *American journal of political science* 49(3):593–608.

Huggler, Justin and Richard Burgess. 2015. "Germany's 'Islamisation' marches: how the pro- and anti-Pegida rallies measured up." The Telegraph.
URL: *http://www.telegraph.co.uk/news/worldnews/europe/germany/11342953/Germanys-Islamisation-marches-how-the-pro-and-anti-Pegida-rallies-measured-up.html*

Humphreys, Macartan and Jeremy M Weinstein. 2008. "Who fights? The determinants of participation in civil war." *American Journal of Political Science* 52(2):436–455.

Inge, Sophie. 2013. "Nazi graffiti and pig's head left at mosque site." The Local.
URL: *http://www.thelocal.fr/20130415/nazi-graffiti-and-pigs-head-found-on-mosque-site*

Isaac, Mike. 2016. "Twitter Steps Up Efforts to Thwart Terrorists' Tweets." New York Times.
URL: *http://www.nytimes.com/2016/02/06/technology/twitter-account-suspensions-terrorism.html*

Jacobs, Justin. 2014. "Criminal Complaint: Nicholas Michael Teausant." United States District Court for the Eastern District of California.
URL: *https://www.justice.gov/sites/default/files/usao-edca/legacy/2014/03/17/Teausant%20Complaint%2*

James, Gareth, Daniela Witten, Trevor Hastie and Robert Tibshirani. 2013. *An introduction to statistical learning.* Vol. 112 Springer.

Jurgens, David. 2013. "That's What Friends Are For: Inferring Location in Online Social Media Platforms Based on Social Relationships." *ICWSM* 13:273–282.

Jurgens, David, Tyler Finethy, James McCorriston, Yi Tian Xu and Derek Ruths. 2015. Geolocation Prediction in Twitter Using Social Networks: A Critical Analysis and Review of Current Practice. In *Proceedings of the 9th International AAAI Conference on Weblogs and Social Media (ICWSM).*

Kaushal, Neeraj, Robert Kaestner and Cordelia Reimers. 2007. "Labor market effects of September 11th on Arab and Muslim residents of the United States." *Journal of Human Resources* 42(2):275–308.

Kent Online. 2015. "MP Gareth Johnson contacts police over Britain First "anti-halal operation" in Dartford." Kend Online.
URL: *http://www.kentonline.co.uk/dartford/news/mp-contacts-police-over-far-right-38953/*

Kibris, Arzu. 2011. "Funerals and elections: The effects of terrorism on voting behavior in Turkey." *Journal of Conflict Resolution* 55(2):220–247.

King, Brian. 2016. "Criminal Complaint: Gregory Hubbard, Dayne Antani Christian, and Darren Arness Jackson." United States District Court for the Southern District of Florida.
URL: *https://www.justice.gov/opa/file/878546/download*

King, Gary and Langche Zeng. 2001. "Logistic regression in rare events data." *Political analysis* 9(2):137–163.

Klausen, Jytte. 2015. "Tweeting the jihad: Social media networks of Western foreign fighters in Syria and Iraq." *Studies in Conflict & Terrorism* 38(1):1–22.

Kydd, Andrew H and Barbara F Walter. 2006. "The strategies of terrorism." *International Security* 31(1):49–80.

Lake, David A. 2002. "Rational extremism: Understanding terrorism in the twenty-first century." *Dialogue IO* 1(01):15–29.

Last, John M, Robert A Spasoff, Susan S Harris and Michel C Thuriaux. 2001. *A dictionary of epidemiology*. International Epidemiological Association, Inc.

Lubbers, Marcel and Peer Scheepers. 2002. "French Front National voting: a micro and macro perspective." *Ethnic and Racial Studies* 25(1):120–149.

Malm, Sara. 2015. "Muslim is stabbed to death while shopping in Dresden, where thousands have gathered for right wing rallies." Daily Mail.
URL: *http://www.dailymail.co.uk/news/article-2909950/PEGIDAs-anti-Muslim-calls-shake-German-politics.html*

Marans, Daniel. 2015. "What You Need To Know About Syrian Refugees Who Have 'Positive' Views Of ISIS." Thw World Post.
URL: *http://www.huffingtonpost.com/entry/syrian-refugees-positive-views-isis_us_564e2c72e4b08c74b734fc83*

Marcus, George E. 1995. *With malice toward some: How people make civil liberties judgments*. Cambridge University Press.

McCauley, Clark and Sophia Moskalenko. 2008. "Mechanisms of political radicalization: Pathways toward terrorism." *Terrorism and political violence* 20(3):415–433.

McGee, Jeffrey, James A Caverlee and Zhiyuan Cheng. 2011. A geographic study of tie strength in social media. In *Proceedings of the 20th ACM international conference on Information and knowledge management*. ACM pp. 2333–2336.

McGee, Jeffrey, James Caverlee and Zhiyuan Cheng. 2013. Location prediction in social media based on tie strength. In *Proceedings of the 22nd ACM international conference on Conference on information & knowledge management*. ACM pp. 459–468.

Merari, Ariel. 1993. "Terrorism as a Strategy of Insurgency." *Terrorism and political violence* 5(4):213–251.

Meyer, Henning and Ulrich Storck. 2015. "Understanding Pegida – An Introduction.".
URL: *https://www.socialeurope.eu/wp-content/uploads/2015/03/Pegida-eBook.pdf*

Montalvo, Jose G. 2011. "Voting after the bombings: A natural experiment on the effect of terrorist attacks on democratic elections." *Review of Economics and Statistics* 93(4):1146–1154.

Neumann, Peter R. 2013. "The trouble with radicalization." *International Affairs* 89(4):873–893.

Norris, Pippa. 2005. *Radical right: Voters and parties in the electoral market*. Cambridge University Press.

Obama, Barack. 2011. *Empowering local partners to prevent violent extremism in the United States*. Office of the President.

Office for Civil Rights and Civil Liberties. 2016. *Community Engagement Section*.
URL: *https://www.dhs.gov/sites/default/files/publications/community-engagement-dhs-crcl-poster.pdf*

Office for National Statistics. 2016. "Super Output Area (SOA).".
URL: *http://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/guide-method/geography/beginner-s-guide/census/super-output-areas–soas-/index.html*

Panagopoulos, Costas. 2006. "The polls-trends Arab and Muslim Americans and Islam in the aftermath of 9/11." *Public Opinion Quarterly* 70(4):608–624.

Patall, Erika and Harris Cooper. 2008. "Conducting a Meta-Analysis." *The SAGE handbook of social research methods. Los Angeles: SAGE* .

Patel, Faize and Michael German. 2015. "Countering Violent Extremism: Myths and Fact.".
URL: *https://www.brennancenter.org/sites/default/files/analysis/102915%20Final%20CVE%20Fact%20Sh*

Petersen, Roger D. 2001. *Resistance and rebellion: lessons from Eastern Europe*. Cambridge University Press.

Peterson, Robert A and Steven P Brown. 2005. "On the use of beta coefficients in meta-analysis." *Journal of Applied Psychology* 90(1):175.

Pooley, Elizabeth. 2015. A New Sisterhood: The Allure of ISIS in Syria For Young Muslim Women in the UK PhD thesis Arizona State University.

Quillian, Lincoln. 1995. "Prejudice as a response to perceived group threat: Population composition and anti-immigrant and racial prejudice in Europe." *American sociological review* pp. 586–611.

Rabasa, Angel, Stacie L Pettyjohn, Jeremy J Ghez and Christopher Boucek. 2010. "Deradicalizing Islamist Extremists." *DTIC Document* .

Radio Free Europe Radio Liberty. 2016. "Prague Refugee Center Fire-Bombed After Anti-Muslim Protests.".
URL: *http://www.rferl.org/content/pegida-refugee-crisis-protests-prague-attack/27536604.html*

Raghavan, Sudarsan. 2016. "Two of her daughters joined ISIS. Now she's trying to save her two younger girls." The Washington Post.
URL: *https://www.washingtonpost.com/world/middle_east/two-of-her-daughters-joined-isis-now-shes-trying-to-save-her-two-younger-girls/2016/05/25/700ea478-1d1e-11e6-82c2-a7dcb313287d_story.html*

Ramos, Juan. 2003. Using tf-idf to determine word relevance in document queries. In *Proceedings of the first instructional conference on machine learning*.

Reuters. 2016. "Anti-Islam movement PEGIDA stages protests across Europe." Reuters.
URL: *http://www.reuters.com/article/us-europe-migrants-protests-germany-idUSKCN0VF0P4*

Revkin, Mara and Ahmad Mhidi. 2016. "Why ISIS Is Rooting for Trump." *Foreign Affaird* .

Robinson, Belinda. 2015. "ISIS supporter, 17, is sentenced to 11 years in prison for tweeting 7,000 messages backing the terror group and helping a friend to join fighters in Syria." Daily Mail.
URL: *http://www.dailymail.co.uk/news/article-3214436/Virginia-teenager-ran-pro-ISIS-Twitter-account-sentenced-11-years-prison.html*

Romaniuk, Peter. 2015. *Does CVE Work? Lessons Learned from the Global Effort to Counter Violent Extremism*. Global Center on Cooperative Security.

Rosendorff, B Peter and Todd Sandler. 2010. "Suicide terrorism and the backlash effect." *Defence and Peace Economics* 21(5-6):443–457.

RT News. 2016. "Dozens arrested as Pegida anti-migrant marches sweep across Europe." RT News.
URL: *https://www.rt.com/news/331586-pegida-anti-migrant-demonstrations-europe/*

Rydgren, Jens. 2008. "Immigration sceptics, xenophobes or racists? Radical right-wing voting in six West European countries." *European Journal of Political Research* 47(6):737–765.

Savage, Timothy M. 2004. "Europe and Islam: Crescent waxing, cultures clashing." *Washington Quarterly* 27(3):25–50.

Scacco, Alexandra. 2008. "Who riots? Explaining individual participation in ethnic violence." *Unpublished manuscript, New York University* .

Schlanger, Margo. 2011. *Office for Civil Rights and Civil Liberties (CRCL) Newsletter*. Vol. 2 U.S. Deparment of Homeland Security.
URL: *https://www.dhs.gov/sites/default/files/publications/crcl-newsletter-september-2011_0.pdf*

Schmitt, Eric and Somini Sengupta. 2015. "Thousands Enter Syria to Join ISIS Despite Global Efforts." New York Times.
URL: *http://www.nytimes.com/2015/09/27/world/middleeast/thousands-enter-syria-to-join-isis-despite-global-efforts.html*

Sedgwick, Mark. 2010. "The concept of radicalization as a source of confusion." *Terrorism and Political Violence* 22(4):479–494.

Shane, Scott, Matt Apuzzo and Eric Schmitt. 2015. "Americans Attracted to ISIS Find an 'Echo Chamber' on Social Media." International New York Times.
URL: *http://www.nytimes.com/2015/12/09/us/americans-attracted-to-isis-find-an-echo-chamber-on-social-media.html?_r=0*

Sides, John and Kimberly Gross. 2013. "Stereotypes of Muslims and Support for the War on Terror." *The Journal of Politics* 75(03):583–598.

Stack, Liam. 2015. "American Muslims Under Attack." New York Times.
URL: *http://www.nytimes.com/interactive/2015/12/22/us/Crimes-Against-Muslim-Americans.html?_r=0*

Stephan, Walter G and Cookie W Stephan. 1996. *Intergroup relations.* Brown & Benchmark Publishers.

Sullivan, John L, James Piereson and George E Marcus. 1993. *Political tolerance and American democracy.* University of Chicago Press.

Sykes, Selina. 2016. "Far-right support surges in Belgium as anger grows over Brussels terror attacks." Express.
URL: *http://www.express.co.uk/news/world/655822/Brussels-attacks-Belgium-far-right-Flemish-parties-nationalist-Flanders-independence*

Tajfel, Henri and John C Turner. 1979. "An integrative theory of intergroup conflict." *The social psychology of intergroup relations* 33(47):74.

Takhteyev, Yuri, Anatoliy Gruzd and Barry Wellman. 2012. "Geography of Twitter networks." *Social networks* 34(1):73–81.

Theocharis, Yannis, Pablo Barberá, Zoltán Fazekas, Sebastian Adrian Popa and Olivier Parnet. 2016. "A Bad Workman Blames His Tweets: The Consequences of Citizens' Uncivil Twitter Use When Interacting With Party Candidates." *Journal of Communication* 66(6):1007–1031.

The Telegraph. 2016. "Thousands march in anti-Islam Pegida rallies across Europe." The Telegraph.
URL: *http://www.telegraph.co.uk/news/worldnews/europe/germany/12144357/Thousands-march-in-anti-Islam-Pegida-rallies-across-Europe.html*

Todd, Tony. 2015. "French far-right sees popularity rise after Paris attacks." France 24.
URL: *http://www.france24.com/en/20151123-french-far-right-popularity-attacks-paris*

U.S. Department of Homeland Security. 2016. *Office for Civil Rights and Civil Liberties.*
URL: *https://www.dhs.gov/office-civil-rights-and-civil-liberties*

Van Zeller, Mariana. 2016. "Radicals Rising." Fusion.
URL: *http://interactive.fusion.net/radicals-rising/*

Vanden Brook, Tom. 2015. "ISIL activity drives up Pentagon threat level.".
URL: *http://www.usatoday.com/story/news/nation/2015/05/08/pentagon-security-isis/26976725/*

Velasco González, Karina, Maykel Verkuyten, Jeroen Weesie and Edwin Poppe. 2008. "Prejudice towards Muslims in the Netherlands: Testing integrated threat theory." *British Journal of Social Psychology* 47(4):667–685.

Vidino, Lorenzo and Seamus Hughes. 2015*a*. "Countering Violent Extremism in America.".

Vidino, Lorenzo and Seamus Hughes. 2015*b*. "ISIS in America: From Retweets to Raqqa (George Washington University: Program on Extremism, 2015) 7.".

Viechtbauer, Wolfgang et al. 2010. "Conducting meta-analyses in R with the metafor package." *J Stat Softw* 36(3):1–48.

Wike, Richard and Brian J Grim. 2010. "Western views toward Muslims: Evidence from a 2006 cross-national survey." *International journal of public opinion research* p. edq002.

Wiktorowicz, Quintan. 2005. *Radical Islam rising: Muslim extremism in the West*. Rowman & Littlefield Publishers.

Wilner, Alex S and Claire-Jehanne Dubouloz. 2010. "Homegrown terrorism and transformative learning: an interdisciplinary approach to understanding radicalization." *Global Change, Peace & Security* 22(1):33–51.

Wilson, Lydia. 2015. "What I Discovered From Interviewing Imprisoned ISIS Fighters." The Nation.
URL: *http://www.thenation.com/article/what-i-discovered-from-interviewing-isis-prisoners/*