Title
Essays on Protest Mobilization in Authoritarian Regimes

Author
Steinert-Threlkeld, Zachary

Publication Date
2016

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Essays on Protest Mobilization in Authoritarian Regimes

A Dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Political Science

by

Zachary Caleb Steinert-Threlkeld

Committee in charge:

Professor Emile M. Hafner-Burton, Chair
Professor James H. Fowler, Co-Chair
Professor J. Lawrence Broz
Professor David A. Lake
Professor Barbara F. Walter

2016

The Dissertation of Zachary Caleb Steinert-Threlkeld is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Co-Chair

Chair

University of California, San Diego

2016

DEDICATION

To my parents, Kayte and Tom.

EPIGRAPH

Disobedience is the true foundation of liberty.

— Henry David Thoreau

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

Undergraduates are supposed to make lifelong friends with their freshman floor, a random assortment of people one should enjoy simply because of proximity. I did not. I have, however, established lifelong bonds with my Ph.D. peers, a random assortment of people one should enjoy simply because of proximity. As promised, the previous 6 years tried and exhilarated me, but I would not trade them for anything else. To everyone who has contributed to that feeling, thank you.

It turns out that the Ph.D. experience is determined as much by, if not more than, one's peers as it is any other factor. Nathan Combes and Blake McMahon have provided endless feedback and encouragement, not to mention great friendship. Maya Duru, Scott Guenther, and Aditya Ranganath were part of a third year prospectus group that kept me focused and moving forward. (Scott's constant beating of me in basketball has also taught me that I should stick to racquet sports, though the three times I have bested him have been three of my happiest days in San Diego.) Alex Verink shares my coffee addiction and interest in communications technology; his presence as an officemate and neighbor has been nothing but positive. Konstantin Ash, Will Hobbs, and Sara Kerosky should all receive their own sentences, and that they do not reveals why one should not write an acknowledgment section last. Others who have helped me, professionally and personally, include: Kate Blackwell, Rob Bond, Garrett Bredell, Chris Chiego, Aaron Cotkin, Cesi Cruz, Michael and Emilie Davidson, Rex Douglass, Chris Fariss, Dotan Haim, Alex Hughes, Ryan Jablonski, Andy Janusz, Jason Jones, Jeff Kaplow, David Lindsey, Adeline Lo, Yon Lupu, Dan Maliniak, Rupal Mehta, Mike Nicholson, Nick Obradovich, Kai Ostwald, Mike Rivera, Alex Ruiz-Euler, Jaime Settle, Yunkyu Sohn, David Straszheim, Clara Suong, Mona

Vakilifathi, Matt and Mandi Walsh, Louis Weimer, Jason Wu, and Stephanie Young.

As an undergraduate, Robert Canfield warned me not to pursue a Ph.D. because the job market is horrendous; I learned therefore to apply only to top programs, and I will be forever grateful for his admonition. Equal thanks go to my thesis advisor, Bret Gustafson; then a young father up for tenure, he found time to be gracious with his feedback. Tim Parsons led a freshman development seminar that challenged and stimulated me, leaving me wanting more; his warning that my attention resembled a cow at pasture continues to motivate me to stay focused.

I could not have expected to meet so many great minds. James Fowler and Emilie Hafner-Burton have shown how to be a scholar, mentor, and maintain a personal life. One improves by constantly working outside of a comfort zone, and their support, guidance, and occasional tough-love made me much stronger than when I started. Equal thanks go to Lawrence Broz, for reading closely everything I sent; David Lake, for long office chats and focus on theory development; and Barbara Walter, for forcing me to focus on theory and understand how to succeed professionally. The International Relations faculty as a whole deserve special thanks. Jesse Driscoll, Erik Gartzke, Peter Gourevitch, Steph Haggard, Megumi Naoi, Phil Roeder, Christina Schneider, Branislav Slantchev, and David Victor have all contributed to my growth in incalculable ways.

I would also be remiss not to thank the political science staff; they are the ones who make sure dozens of disorganized professors and even more disorganized graduate students functions at a reasonable level. Steve Carlson deserves special thanks for running the Laboratory on International Law and Regulation and knowing cool places to which to travel. Elizabeth Brost, Christina Butler, James Fritsch, Brian

Hile, Collette Isachsen, Jennifer Neri, Ariane Parkes, Arturo Vasquez, Regina Ready, and Natacha Scherf have been as important as anyone else in helping me enjoy my time in San Diego.

Even better than completing the Ph.D. has been meeting Jessica Jinn. I was in my second year, she in her first at the School of Global Policy Studies, and we immediately bonded. A girl who reads *The Economist*, brews Intelligentsia coffee, and believed as a kid that she had to marry a Jewish man to get to heaven - what more could a guy want? She has persevered through my successes and failures and understood when I needed to ignore her to focus on work. We marry one week after graduation, and I could not ask for anyone better with whom to start adulthood.

I would not have been in the position to succeed in college and graduate school without my mom, dad, and brother. From putting me on an allowance to minimizing my college debt, proofreading middle school essays to ensuring I always had reading material, and letting me go to nerd summer camp, they ensured I would look back to childhood with only positive thoughts. (Well, except for that overreaction regarding a speeding ticket.) My brother has been my best friend since I consistently beat him at ping pong; once he started beating me, just a few years ago, his laid-back intelligence and fraternal banter ensured he stayed in my good graces.

In an era of diminishing support for the public sector, I would like to call special attention to how I have survived the previous 6 years. There are few American public universities that are world-class, and the majority of those are part of the University of California system. Because of California's far-sighted support of higher education, I have been privileged to receive a livable, but not overly generous, salary and medical plan, and my professors, all world-class intellects, are justly compen-

sated. Undergraduates, many of whom are first generation college students, receive an education on par with the best private universities and pay, though more than they should, much less than they would at those places. Concurrent with a community college system that allows students to transfer into the University of California, there are few institutions as inspirational as the University of California. In addition, I have received grant and fellowship support from the United States Agency for International Development and the Institute on Global Conflict and Cooperation, both public institutions, and I will soon start as a professor at the University of California, Los Angeles. Except for four years as an undergraduate, I am a product of public education and proof that it should be defended, if not strengthened.

Determining where, when, and against whom to draw the line for acknowledgement is a difficult, arbitrary decision. Exclusion from here is not exclusion from my appreciation.

*****

Chapter 1, in part, has been submitted for publication. I am the primary investigator and author of this paper.

Chapter 2, in part, has been submitted for publication. I am the primary investigator and author of this paper. This chapter is made possible with support from the American people through the United States Agency for International Development (USAID) and under the Democracy Fellows and Grants (DFG) program implemented by IIE. The contents are the sole responsibility of the author and do not necessarily reflect the views of USAID, the United States Government, or IIE.

Chapter 3, in part, has been submitted for publication. I am the primary investigator and author of this paper. This chapter is made possible with support from the American people through the United States Agency for International Development (USAID) and under the Democracy Fellows and Grants (DFG) program implemented by IIE. The contents are the sole responsibility of the author and do not necessarily reflect the views of USAID, the United States Government, or IIE.

Chapter 4, in part, has been submitted for publication. I am the primary investigator and author of this paper.

VITA

2008	B. A. in Anthropology and Economics *(summa cum laude)*, Washington University in St. Louis
2012	M. A. in Political Science, University of California, San Diego
2016	Ph. D. in Political Science, University of California, San Diego

PUBLICATIONS

Steinert-Threlkeld, Zachary C., Delia Mocanu, Alessandro Vespignani, and James Fowler. 2015. "Online Social Networks and Offline Protest." *EPJ Data Science* 4(19).

ABSTRACT OF THE DISSERTATION

Essays on Protest Mobilization in Authoritarian Regimes

by

Zachary Caleb Steinert-Threlkeld

Doctor of Philosophy in Political Science

University of California, San Diego, 2016

Professor Emile M. Hafner-Burton, Chair
Professor James H. Fowler, Co-Chair

How do individuals in non-democracies organize collective action? Throughout history, it is common for those in power to structure society so as to make mass action highly unlikely; yet, from peasant revolts in the Middle Ages to urban arisings during the Arab Spring, individuals surmount these barriers. This dissertation argues that they surmount these barriers because mobilization occurs through interpersonal connections that connect individuals. These social networks transmit information necessary for protest mobilization, while state repression is most effective against institutions and prominent individuals.

This argument is tested in the context of the Arab Spring. In Chapter 1, 13.8 million geolocated messages and daily data on protests are gathered for the Middle East and North Africa from November 1st, 2010 through the end of 2011. These data reveal that protest mobilization correlates with coordination from individuals on the periphery of their country's social network, not those who are prominent. Because the identification strategy for who is prominent in a network may average away important differences, Chapter 2 focuses on activists in Bahrain and Egypt. A supervised topic model on those activists' messages still shows no effect for coordination; this finding reinforces Chapter 1 since these individuals are those most likely to affect protest mobilization. Qualitative evidence suggests that their effect occurs through generating common knowledge about mass dissatisfaction before protests occur.

Chapters 3 and 4 focus on methodological points. Chapter 3 introduces a technique for using Twitter data to analyze network structure as it evolves. Analyzing network structure over time requires data on the entire network, which is difficult for two reasons. First, Twitter imposes limits on acquiring data that make it effectively impossible to get complete network data. Second, those same restrictions make it effectively impossible to analyze that network as it changes. This chapter introduces a measure that can be collected rapidly enough from Twitter to obviate these issues. Chapter 4 is a general purpose introduction to Twitter. It explains what questions Twitter data can answer, how to acquire those data, and programming approaches.

Chapter 1

Spontaneous Collective Action: Peripheral Mobilization During the Arab Spring

1.1 Abstract

Who is responsible for protest mobilization? Models of disease and information diffusion suggest that those central to a social network (the core) should have a greater ability to mobilize others than those who are less well-connected. To the contrary, this paper argues that those not central to a network (the periphery) can generate collective action, especially in the context of large-scale protests in authoritarian regimes. To show that those on the edge of a social network have no effect on levels of protest, this paper develops a dataset of daily protests across 16 countries in the Middle East and North Africa over 14 months from 2010 through 2011. It combines that dataset with geocoded, individual-level communication from the same period and measures the number of connections of each person. Those on the periphery are shown to be responsible for changing levels of protest, with some evidence suggesting that the core's mobilization efforts lead to fewer protests. These results have implications for a wide range of social choices that rely on interdependent decision making.

1.2 Introduction

Large groups of people acting without centralized leadership can organize protests. Protests occur as a result of decentralized coordination of individuals, and this coordination helps explain fluctuating levels of protest. Individuals in the core of a social network - those such as activists, members of the media, or civil society organizations - do not mobilize protests. Instead, those on the periphery of the network communicate with each other about the near future (where and when to protest) as well as events as they unfold (the presence of police, what the police are doing, supplies needed, and so on). While those at the center of the network do engage in the same behavior as others, their effect is washed out in comparison to that of the masses they try to lead. I call the ability of the periphery to mobilize *spontaneous collective action.*

There exist two competing explanations for how individuals decide to undertake action. Whether deciding to vote [Downs, 1957, Riker & Ordeshook, 1968, Quattrone & Tversky, 1988], join a political organization [McAdam, 1986, Klandermans & Oegema, 1987, González-Bailón et al., 2011], or protest [Moore, 1995, Lichbach, 1998, Goldstone, 2001], individuals may decide to do so as a result of effort from centralized, well-connected individuals (the core) or those on the periphery. Those at the center of a social network can provide focal points for action, alternative policies for voters, new information about policies, or demonstrate a regime is weaker than previously thought, all contributing to individuals taking collective action [Taylor, 1988, Shachar & Nalebuff, 1999, Dalton et al., 2002, Gerber, Karlan & Bergan, 2006]. On the other hand, individuals can decide to vote (or protest or join a movement) based on the influence of those they know [Schussman & Soule, 2005, Gerber, Green & Larimer, 2008], beliefs in their own ability to affect an outcome [Finkel, Muller & Opp, 1989, Goldstone, 1994, Opp, 2012], or from observing the behavior of others [Granovetter, 1978, Lohmann, 1994]. These others are the peripheral members.

This argument is tested using data from the Arab Spring, the protests which started in Tunisia in December 2010 and soon spread through North Africa and the Middle East. The events of the Arab Spring, the most prominent large-scale, widespread protests since the collapse of the Soviet Union, provide an ideal situation in which to test this theory. "Arab Spring" refers to the series of protests which started in Tunisia on December 14^{th}, 2010 (leading to the resignation of that country's president), slowly spread to neighboring countries over the following 6 weeks, and inspired massive turnout in Egypt that caused President Hosni Mubarak to resign on February 11^{th}, 2011. This paper will show that these protests were not driven by the people who had tried for years to organize them. Instead, they were organized by large groups of individuals discussing amongst themselves where to go, how to get there, when to go, and what was going on once there. This paper does not seek to explain the Arab Spring, but it does, in the course of developing the spontaneous collective action theory, present the first large-scale, systematic evidence on how individuals behaved in each country.

To test the core versus peripheral hypotheses, this paper connects two large-scale datasets. First, a machine-coded events dataset, the Integrated Conflict Early Warning System (ICEWS) is combed to measure the number of protests per day across 16 countries from November 1^{st}, 2010 through December 31^{st}, 2011. Second, a dataset of geolocated tweets in the same countries from the same period is built. These 13,754,998 tweets show what was being said, when it was being said, where, and how many connections each tweet author had. Combining these datasets and using a wide range of models and operationalizations, mass mobilization is shown to occur through peripheral individuals.

This paper proceeds in 8 sections. Section 1.3 compares and contrasts existing theories of protest mobilization with the one developed here. In Sections 1.4-1.6, the empirical strategy is discussed. Section 1.7 presents the main findings and a battery of robustness checks to reinforce them. Section 1.8 undertakes a closer look at Egypt to identify core actors, and Section 1.9 concludes with final thoughts and suggestions for future research.

1.3 Theory

Coordination drives protest mobilization, and peripheral members of a network drive coordination. Coordination consists of two components.

First, individuals deciding to protest must receive a credible signal that large numbers of people are protesting, suggesting that the cost of protesting is low. Second, once protests have commenced, information about upcoming protests - when a protest will occur, where protesters will convene, routes to take, supplies needed, and so on - needs to be provided. Members on the periphery of a social network better provide both components of coordination than those in the core. Individuals at the core of a network - those connected to many more people than the median person - are socially distant from most of those connections and few in number. This distance attenuates the weight of the signal the core sends [Centola & Macy, 2007, pgs. 725-726], while their rarity limits the influence of their action and relevance of the protest information they provide.

Networks provide a framework for understanding how a phenomenon spreads between items; when these items are people, the network is a social network, and connections represent two people between whom a phenomenon can spread. These phenomena fall into two categories, simple and complex. A simple contagion is a phenomena which can spread between individuals after one exposure, such as illness or information about job opportunities [Granovetter, 1973]. Disease transmission or news are canonical examples: John only needs to meet one person with the flu to catch it, and Jane only needs to talk with one person to learn tomorrow's weather. John does not become more sick from meeting a second infected person, and Jane does not become more knowledgeable receiving the same weather report from a second person. Except for rare cases, simple contagions always spread in a network [Newman, 2003]. Simple contagions spread quickest when core nodes are affected since those nodes can spread the phenomena in question to many nodes at once, regardless of the structure of the underlying network [Watts, 2004, pgs. 257-260]. In simple contagion models, diffusion of a phenomena is less likely when the diffusion starts on the periphery.

Complex phenomena are those whose transmission requires an individual to observe that phenomena in two or more people. Contact with two or more sources is required when the phenomena possesses positive externalities, gains credibility or legitimacy when multiple people partake, or have an emotional component [Centola & Macy, 2007, pgs. 707-708]. Models of complex contagion are often called threshold models since they require an individual to be exposed to a defined amount of other people in the network before switching states [Granovetter, 1978, Schelling, 1978].[1] The existence of thresholds makes the spread of complex contagions less certain, as network structure - the distribution of thresholds - can cause a contagion to stay trapped in one part of the network [Watts, 2002]. Because contact with more than one source is required for complex contagion's spread, core members do not automatically lead to the spread of the phenomenon in question. The existence of that phenomena in peripheral parts of the network becomes essential for its spread throughout the network.

[1] The threshold is sometimes defined as a constant and sometimes as a fraction of network size. This distinction matters for small networks but large ones. For example, in a network of 8 individuals, a threshold of 1/8 does not represent complex contagion because an individual will switch states when only 1 person it knows has has; in a network of 800,000, a threshold of 1/8 would correspond to a late mover. Because mass protest involves large groups of people, the difference between numeric and proportional thresholds is moot. See Centola and Macy 2007 for an extended discussion on the difference between fractional and numeric thresholds.

Protests are a complex contagion phenomenon because increasing participation makes others more likely to join. Individuals are more likely to protest as others protest, since the cost of protesting decreases as a function of group size. Individuals are especially more likely to protest when they know others who are protesting [Opp & Gern, 1993], and those on the periphery of a network are more likely to know others on the periphery than in the core [McPherson, Smith-Lovin & Cook, 2001, Kwak et al., 2010]. Since there exist many more individuals on the periphery of a network than in the core, protest is therefore more likely to occur when those on the periphery of a network mobilize.

The first mechanism through which the periphery of a social network mobilizes protest is through providing a credible signal about participation in the protest. Peripheral members mobilize other participants better than those in the core because they provide a more credible signal that the protest enjoys widespread participation. If a protest is dominated by core members, the signal suggests that the policy disagreement does not affect many people who do not usually protest. This insight is similar to that made by Susanne Lohmann: she argues that unexpected participation of "moderate activists" drives protest mobilization because "extreme activists" always protest, so their participation is not a credible signal about the severity of a grievance [Lohmann, 1994]. A larger than expected turnout of "moderate activists" signals to others that grievances are widely shared, leading to the expectation that one's action will decisively lead to a policy change. In discussing the effect of network structure on collective action, David Siegel explains that:

> [...] the people at the bottom of the network – the proletariat, if you will - can [mobilize] if they have enough connections among themselves. The key here is to obtain a sufficiently large and well-connected group of people at the bottom of the hierarchy who [...] are highly internally

> motivated to participate. If these requirements are achieved, the bottom of the hierarchy can spur the network on to very high levels of participation. [Siegel, 2009, pgs. 134-135]

Peripheral mobilization dominates core mobilization because there are more people on the periphery.

The second mechanism through which the periphery of a social network mobilizes protest is by providing more information about a protest as it unfolds, and this information has the effect of coordinating protestor movement and tactics. One type of information is situational awareness, knowledge about unfolding events, and peripheral members, because of their number, provide this awareness in ways the core cannot. Situational awareness entails knowing the size of the police presence, which routes police block, whether or not police engage with protesters, paths around police, and where other protest groups find themselves. protesters are also more likely to reach and hold onto their desired site if they can approach it from multiple directions and coordinate their action, as doing so makes it harder for police to contain the protesters [Gunning & Baron, 2013, 168-174]. But, since there exists a finite supply of core individuals, splitting a protest into subcomponents means the ability of core members, who are few in number, to control them is lessened.

Moreover, once engagement with state forces commences, order often dissolves; a protest is a quickly shifting series of actions occurring in an area too big to be observed by a few individuals. During a street engagement with government forces, protesters may require reinforcements on some streets and not others, while supplies such as gas masks necessary in one place but not elsewhere. If a group is able to cause police to retreat, communicating that advantage to nearby protesters can provide reinforcements to exploit this development. But relying on core members, who are few in number, to coordinate these reactions decreases the efficacy with which protesters can react to new developments. Situational awareness therefore increases the likelihood of protest success, and situational awareness is increased when information flows from and between as many individuals as possible.

Situational awareness also entails providing logistical support for a protest. During the initial march to a protest site, one key piece of information is what kind of equipment individuals need. Gas masks, onions, and soda mitigate the effect of tear gas, while hammers, slingshots, and shields are necessary if projectiles are to be employed. Individuals also need to know to where these supplies need to be delivered, as some groups of protesters may be marching peacefully while others in different neighborhoods confront the police. The same logic holds once a protest site, such as a city's main square, is established. At this point, the protest site becomes a miniature city; the provision of food, medical supplies, sanitation, communications equipment, and security all require coordination.

A comparison with military tactics clarifies the importance of situational awareness. The chief advantage of German armored divisions at the start of World War II was coordination enabled by new communication technology. Equipped with radios, unit commanders could communicate with their tanks in real-time, maintaining tactical awareness throughout a battle and so allowing them to exploit enemy weaknesses or cover their own [Citino, 2004]. Iraqi forces in the first Gulf War had not learned this lesson: battlefield commands flowed through centralized headquarters in Baghdad, and Coalition forces were able to bomb these facilities, hindering the ability of Iraqi frontline forces to respond to battlefield developments [Press, 2001]. In protests, it is the police who have traditionally had the coordination advantage because of their distributed communication, while protesters have often lacked a similar ability.

Peripheral individuals are better positioned to coordinate than the core. Even in an authoritarian setting, the existence of widespread discontent is often not a surprise. In Tunisia and Egypt, for example, it was widely known that the regimes were unpopular. In Tunisia, oligarchic elites and weak rule of law alienated large segments of society, from students to the working-class, especially outside of Tunis, and desperation suicides were not uncommon events [Breuer, 2012, Al-Zubaidi & Cassel, 2013]. In Egypt, police indiscretion, religious persecution, and economic instability similarly dispirited a majority of the population [Gunning & Baron, 2013, 97-127]. It was well-understood in these countries that dissent was widespread and a minority of a society benefited from current policies at the expense of most others. Widespread, commonly understood dissatisfaction means that latent desires for policy change are known to exist, rendering the task to one of coordinating protest. The periphery then drives mobilization because it signals that disparate, numerous groups of individuals are acting on this discontent.

Signalling and situational awareness allow peripheral members to coordinate their action. For example, a message such as "#jan25 protests will take place all throughout cairo, including shubra, mohendessin, in front of cairo university and on arab league street" issued on the morning January 25^{th}, the first major day of protests in Egypt, provides information about where individuals who want to protest can join others [Idle & Nunns, 2011, pg. 33]. Information less explicitly about coordination can also have a coordinating effect. A large amount of the communication leading up to a protest focuses on supplies needed, how to dress, how to behave towards the police, and the identity of protesters. This communication does not tell people when or where to go, but it helps them estimate levels of support in the population and danger they may face [Lohmann, 1994, Gerbaudo, 2012]. The more people that provide this information, the easier protest coordination becomes.

The importance of signalling and protest information leads to a primary hypothesis that can be broken down into constituent parts.

H1 As coordination from the periphery of a social network increases, more protests should occur.

- **H1a** As the percentage of protesters who are from the periphery of a country's social network increases, more protests should occur.
- **H1b** As more information about protests comes from the periphery of a country's social network, more protests should occur.

H2 As coordination from the core of a social network increases, there should be no change in the number of protests.

The importance of peripheral participation as a signal of broad support is found in the experience of Egyptian mobilization on January 25^{th}. As groups of protesters marched through outlying neighborhoods, they urged onlookers to leave their shops, apartments, and workplaces. Many did, and the protest size snowballed [pg. 51][Cambanis, 2015]. Protesters also emphasized the different parts of society they represented, with particular care taken to recruit outside of the middle-class as well as emphasize independence from the Muslim Brotherhood [Gunning & Baron, 2013, pg. 180]. The initial mobilization therefore included youth, members of football fan clubs, the poor and working class, in addition to individuals who were habitual protesters. Moreover, these habitual protesters were situated in the core of the Egyptian social network and had tried to initially protest on January 18^{th}; only 5 activists protested, reflecting the importance of mobilization from the periphery [Gunning & Baron, 2013, pg. 91].

That peripheral members of a social network provide more information than the core finds support in other settings as well. In a study of diffusion on Facebook, Bakshy et. al (2012) find that weak ties are responsible for most information diffusion because they are more numerous than strong ties (individuals who interact frequently), just as those on the periphery are more numerous than those in the core. Recruitment to Spain's indignados movement, which started less than 4 months after Egypt's first protests, was characterized by individuals' exposure to the same information from different sources [González-Bailón et al., 2011]. Adoption of political attitudes is also increased after exposures from different sources [Romero, Meeder & Kleinberg, 2011], and controlled experiments have confirmed the importance of multiple sources of exposure for changing health attitudes [Centola, 2010]. Complex contagion also drove mobilization processes during the collapse of the Soviet Union [Opp & Gern, 1993] and the American Civil Rights movement [McAdam, 1986], though scholars at the time did not use that language. Finally, this might be the process which drives mass urban mobilizations that bring together disparate groups of individuals who previously did not engage in anti-regime behaviors [Beissinger, 2013, Tufekci, 2014].

That protest is a complex contagion explains why many states have large domestic intelligence apparatuses and fear mass public gatherings. If an individual desiring to protest is concerned that sharing that information will lead to punishment, individuals are less likely to form connections with other individuals. In network terms, there will be fewer bridges between communities, inhibiting the spread of protest mobilization information. (If protest were a simple contagion phenomenon, a few number of protesters could have a large effect, and governments would have to make the costs of protest so high to prevent any display of anti-regime sentiment.) Large public gatherings therefore provide one of the few occasions individuals have of bridging their immediate social communities; these bridges may cause individuals' protest thresholds to be surpassed, and a chain reaction of protests may ensue. For example, protests in Egypt against the Iraq War and marking the third anniversary of the Second Intifada led to the first large-scale public chants against Hosni Mubarak and started the process by which previously disconnected groups of individuals began to coordinate their anti-regime actions [Gunning & Baron, 2013, pgs. 39-47]. In Russia in 1917, an industrial lockout, International Women's Day, and military leave brought together tens of thousands of workers, women, and disgruntled soldiers into the streets of St. Petersburg; the Romanovs fled 4 days later [Kuran, 1989, pg. 63]. China even allows criticism of government officials and policy so long as it does not lead to appeals for collective action [King, Pan & Roberts, 2013].

That protest is a complex contagion phenomenon also does not mean core members are unimportant in terms of protest mobilization. There are at least 3 mechanisms by which core members can facilitate protests: convincing individuals to blame their dissatisfaction on government policies, revealing the state is weaker than commonly believed, and fostering group-identity. First, a core member can help those on the peripheral ascribe their policy dissatisfaction to specific policies of those in power because the information to assign blame is a simple contagion phenomenon. As Javeline summarizes: "individuals faced with any grievance should be more likely to protest if they can make specific attributions of blame for the grievance and that one mechanism by which entrepreneurs [core individuals] might solve collection action problems is by first solving blame attribution problems" [Javeline, 2003, pg. 119]. Second, core members can engage in violence which, if not terminated, reveals that anti-regime preferences and widespread and the regime may be weak [De Mesquita, 2010]. Third, core members can create norms of solidarity, causing individuals to calculate their participation based on group gains [Goldstone, 1994]. Once individuals see themselves as part of a larger group, the benefits of protest increase while the costs decrease, making them more likely to mobilize when the opportunity arises.

These core-based mechanisms are not related to protest mobilization, however, as they occur before mass protests. They predispose individuals to be ready to mobilize, but they do not directly mobilize. In the language of Timur Kuran, they cause preferences to change, but they do not provide the initial spark [Kuran, 1989, pgs. 63-66]. The theory of spontaneous collective action also treats the spark as exogenous.

1.3.1 Scope Conditions

There are two primary scope conditions to the application of the theory to the Arab Spring. First, a country's regime type may determine whether or not protest is a complex contagion phenomenon. Arrests, perfunctory trials, and long jail terms were standard state practices from Morocco to Bahrain [Bellin, 2012, Gunning & Baron, 2013, Khatib & Lust, 2014], making it difficult for some core social network members to organize collective action. Second, mobilization is bounded by the costs a state imposes on protesting. Libya, Saudi Arabia, Bahrain, Syria, and Egypt in 2013 engaged in sustained violent repression of collective action, with heterogenous outcomes.

Authoritarian regimes are likely to repress individuals who impugn them, as targeted repression is a more effective tactic than its indiscriminate killing [Siegel, 2011]. Arbitrary jailing, torture, forced exile, and threats to family are all common tactics used to silence anti-regime individuals. In countries where those who desire policy change and are central to a network are routinely intimidated or silenced, they may not have the ability or desire to engage in coordination activities, and coordination would necessarily occur through those on the periphery. Moreover, in countries tolerant of mass gatherings, individuals may have lower thresholds of participation since they do not fear repression. If an individual does not expect protest to be large to be safe, he or she may join a protest alone or after hearing about it from a core social network member. In these cases, protest is more likely to be a simple contagion event and so be more affected by core members of a social network.

Second, any state can stop protests if it is willing to impose high enough costs [Blaydes & Lo, 2011]. In March of 1988 in Burma, protests started over an event just as random as a fruit vendor lighting himself on fire: a youth arrested for fighting other youth was released from jail through political connections. Tension boiled over the summer, a general strike started on August 8^{th}, and the state engaged in ambiguous amounts of repression. On September 18^{th}, repression became less ambiguous as a result of an army coup; the ensuing repression resulted in at least 1,000 deaths in Rangoon, 3,000 nationwide [Ferrara, 2003]. Protests, which had been stronger throughout August but were tapering by mid-September, ceased. In 1989, a protest movement in China grew over the course of several months; by the end of May, Beijing hosted 250,000 soldiers, and multi-day violent repression began on June 3^{rd}. That repression, in conjunction with the arrest of party leaders and Communist Party reformers, squelched the movement. In 2011 in Egypt, individuals soon realized that the armed forces were not going to repress protests, yet in August 2013, the Egyptian army massacred hundreds, perhaps thousands, of pro-Muslim Brotherhood supports who were staging long-term sit-ins at two Cairo squares after the July 3^{rd} coup against President Mohamed Morsi [Meirowitz & Tucker, 2013]. Repression against secular activists also increased, with those continuing to protest facing lengthy jail sentences or death [Mackey, 2015]. Bahrain's security forces killed protesters at the Pearl Roundabout, after welcoming a coalition of forces from Gulf states; leaders of al-Wefaq, the main Shia opposition party that participation in government before the start of protests, are now in jail, and the party's leader faces a 4 year sentence for inciting violence against the monarchy [Kerr, 2015]. While a state faces internal and external costs from repression, the ultimate success of any protest mobilization depends on the state's willingness to repress.

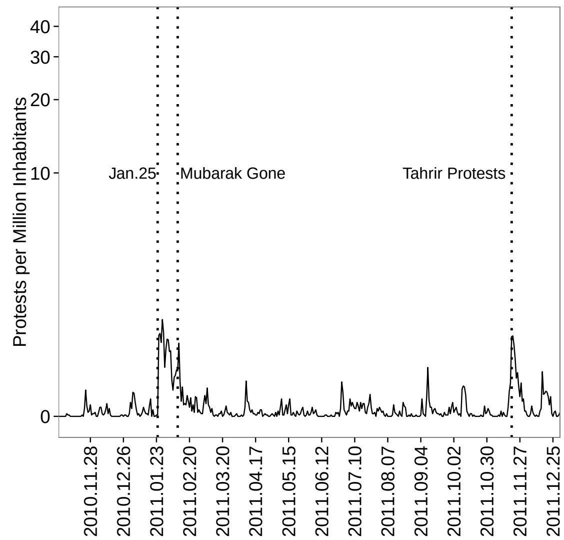
1.4 Data

The Integrated Conflict Early Warning System, a machine-coded events dataset that reads newspaper articles, provides the dependent variable, number of protests across 16 countries in the Middle East and North Africa from November 1^{st}, 2010 through December 31^{st}, 2011[Boschee et al., 2015]. ICEWS codes 20 categories of events of increasing severity, from public statements through unconventional mass violence. All events coded as protests where the event occurs in one of these 16 countries country are kept. These countries are Morocco, Algeria, Tunisia, Libya, Egypt, Lebanon, Syria, Jordan, Saudi Arabia, Oman, Yemen, Bahrain, the United

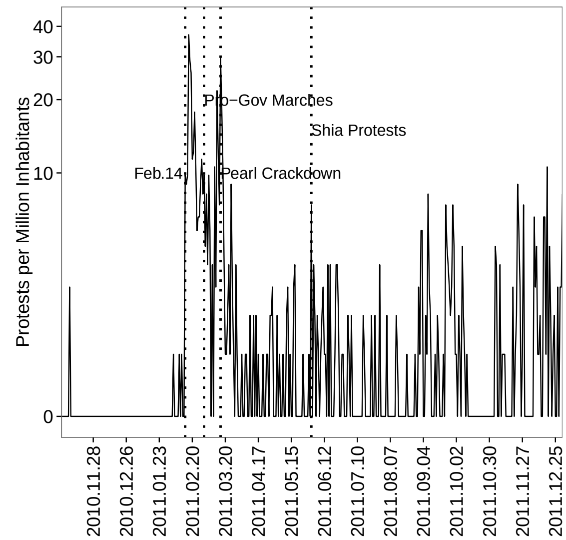
Arab Emirates, Kuwait, Iraq, and Qatar; Israel is excluded. Figure 1.1 shows the ICEWS recording of protest in two high-protest (Egypt, Bahrain) and low-protest (Morocco, Qatar) countries.

Social media data are ideal for understanding protest, for three reasons. First, in states that control information disseminated through newspapers, radio, and television, social media are one of the few independent sources of information [Edmond, 2013]. Social media have therefore become a tool for citizens to gather and disseminate information in information-scarce environments such as authoritarian regimes. In this way, social media may have a similar effect as independent media [Egorov, Guriev & Sonin, 2009] or the disclosure of economic data by an autocrat [Hollyer, Rosendorff & Vreeland, 2015]. Second, state actors belatedly realized the power of social media, leaving it unregulated; lack of regulation made social media an attractive tool for anyone seeking independent information, and the information contained in social media therefore more closely reflected the offline world than did official news sources [Hamdy & Gomaa, 2012]. Social media has therefore become a critical component of many protest movements, starting with the 2009 Iran election protests and continuing through the Ukraine civil war [Burns & Eltham, 2009, Rahimi, 2011]. Third, it provides the best temporal resolution of any data source. It is therefore one of the few sources available to researchers interested in dynamic processes that can provide micro-level information on these processes.

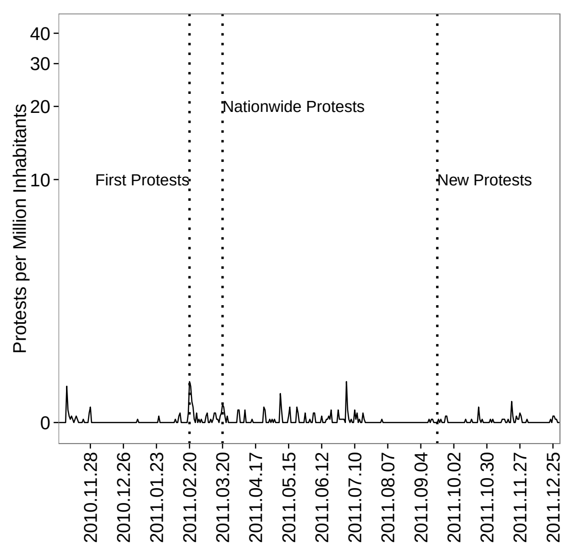
Twitter, a global social media platform, provides data on daily, individual-level communication. It is a global social network with over 500 million users generating almost 500 million daily messages (tweets). Anyone with an internet connection or phone can access it, and most users create and consume content using their mobile



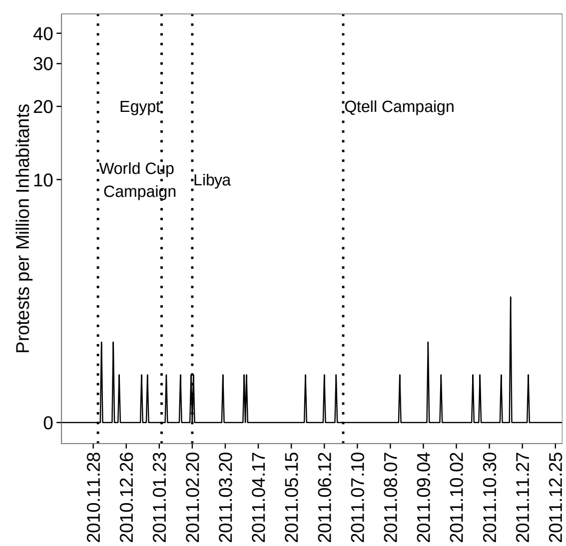
(a) Egypt



(b) Bahrain



(c) Morocco



(d) Qatar

This figure shows that ICEWS captures different levels of intensity of each country's protests, both temporally and in cross-section. Egypt, which experienced sustained, widespread protest, has the most recorded protests of any country in the dataset, but it has fewer per person than Bahrain. Morocco had a sustained protest campaign that did not mobilize as many people as Egypt or Bahrain, and Qatar experienced no protests. ICEWS' count of protests also varies during days they are expected to.

Figure 1.1: Protests per Million Inhabitants

devices; contrary to popular belief, one can compose and consume tweets from any kind of phone, though smartphones greatly facilitate the process. The company does not edit or censor its users' tweets, so the content of the network reflects what individuals are discussing at any moment.[2] Only China and North Korea have completely blocked access to it, though countries have temporarily blocked it at different times.

There are four reasons to prefer Twitter as a data source to other social media platforms. First, it is one of the most used social media platforms, usually second only to Facebook [Duggan & Smith, 2013]. Second, it is often used during crisis events to disseminate information, including during protests [Tonkin, Pfeiffer & Tourte, 2011, Earl et al., 2013]. Third, though it is used to discuss political events such as protests, it is also used to engage in quotidian topics like celebrity gossip, the weather, and sports [Boyd, Golder & Lotan, 2010, Sinha et al., 2013]. In the sample of tweets used later to train the support vector machine, *almost 75% were not about political events.* Fourth, Twitter provides a large amount of its data through two programming interfaces, making Twitter data easier to obtain than Facebook's. While other sites with social networking components, such as YouTube or reddit, are also relatively easy to gather data from, none are also used as comprehensively as Twitter.

Moreover, the norms of communication on Twitter makes this the most reliable way to measure coordination across so many countries and days. There are four ways to a user can modify a plaintext tweet. The most common is the # symbol, known as the hashtag. Individuals will affix a hashtag to the front of a word to associate it with a certain conversation, e.g. "Eyewitnesses: NDP thugs throwing molotov cocktails inside the Egyptian Museum. I repeat NDP thugs, NOT anti-

[2]Twitter will censor tweets to comply with countries' laws. For example, it has censored a neo-Nazi group's tweets in Germany and has started to delete accounts from the Islamic State of Iraq and Syria that are deemed to incite violence.

Mubarak protesters. #Jan25 #fb". If a different user then searches for messages containing "#Jan25" or "#fb", this tweet will be returned; employing a hashtag therefore makes the information in one's message more likely to spread beyond just one's social network [Romero, Meeder & Kleinberg, 2011]. Users quickly converge on a few hashtags to use for an event, whether that event is a protest, sporting event, or pop culture meme [Bruns & Burgess, 2011, Lehmann et al., 2012].

Twitter makes it easy to find all tweets containing a hashtag. A user interested in upcoming protests could therefore search, from her smartphone or a computer, for "#jan25", "#egypt", or other hashtags and retrieve every tweet containing those hashtags.[3] That person is therefore quickly exposed to vastly more information than she could gain from traditional interpersonal communication, and she knows that everyone else searching those hashtags will see the same tweets. She is therefore confident that when she reads about the meeting in Batal Ahmed street, many others have read about it as well, and others who search for "#jan25" know that others have seen that tweet as well. The prevalent use of hashtags, convergence to very few during major events, and ease of finding information related to the hashtag make tweets with hashtags they key coordination mechanism.

Twitter data comes courtesy of researchers at Northeastern University's Laboratory for the Modeling of Biological and Socio-Technical Systems [Mocanu et al., 2013]. The tweets involved in this analysis were extracted from Twitter's 10% API, an unbiased sample of 10% of all public activity on the platform. There are two ways in which country of origin was identified. First, if a Twitter user has enabled location

[3] This search is not case sensitive: a user searching for "#jan25" will see the same results as one searching for "#Jan25". The searched term will not return tweets that use the character string without a hashtag, e.g. a tweet that says "police thugs r everywhere in egypt jan25" will not show up in search results.

sharing, the tweet will have GPS coordinates, and Twitter will assign a 2 letter country code to those tweets. Each tweet is then read for a 2 letter code corresponding to one of the 16 countries and saved if there is a match. Second, users can report their location as part of their profile, and that location is reported as metadata with each tweet. The user-reported location is then compared to a dictionary of cities and country names to assign each tweet to a city or country; if that location is part of this study, the tweet is saved.[4] Unlike previous studies that analyze contentious events, tweets in this dataset were not selected based on hashtags, providing a representative sample of what Twitter users were actually talking about, e.g. protests or the weather, during this period. Only 19.74% of all tweets in this sample contain a hashtag, and most are apolitical.

1.95% of tweets in this sample have GPS coordinates, with the location of information of the rest coming from user-reported location. These numbers correspond with other work that finds more than an order of magnitude more tweets when using self-reported location [Leetaru et al., 2013]. It is worth nothing that tweets in the United States with GPS coordinates exhibit bias towards urban areas, non-whites, and high-income groups [Malik et al., 2015], and there is some evidence that users of Twitter in Egypt tend to be well-off individuals in cities [Tufekci & Wilson, 2012]. Malik et. al (2015) do not include tweets with user-reported location and Tufekci and Wilson (2012) do not ask whether users geotag their tweets, so it is unclear if using user-reported location removes these biases.

[4]For more detail, see the Materials and Methods section of Mocanu et. al (2013). That paper uses only tweets from that stream with GPS coordinates for its analysis, whereas I use tweets with GPS coordinates or user-reported location because there were not enough tweets with GPS coordinates in the countries in this study.

1.5 Measures

1.5.1 Coordination

A Gini coefficient for hashtags operationalizes coordination. The Gini coefficient, which ranges from 0 to 1, usually measures income inequality, but it can be used on any orderable discrete quantity. Instead of measuring wealth per person, it here measures occurrences per hashtag per day per country; a 0 means every observed hashtag occurs the same number of times, 1 that one hashtag accounts for all hashtags used in that country on that day. This measure is labeled $Coordination_{i,t}$ for the rest of the paper, and the Supplementary Materials provide a graphical explanation of the operationalization.

Equation 1.1 shows this calculation. For each day t in each country i, there exist n unique hashtags. $Coordination_{i,t}$ counts the number of times each hashtag j occurs (c_j) and uses those counts to calculate the Lorenz Curve of (hashtag) inequality.

$$Coordination_{i,t} = \sum_{i=1}^{16} \sum_{t=1}^{426} \frac{2 \sum_{j=1}^{n} j c_j}{n \sum_{j=1}^{n} c_j} - \frac{n+1}{n} \tag{1.1}$$

Three other Twitter behaviors that may impact coordination are measured. Retweets, equivalent to forwarding an e-mail to one's entire contact list, can also promote coordination An example of a retweet is, "RT @Ekramibrahim: Police, specially in civil clothes are holding electricity sticks. #jan25". Ekramibrahim is the author of the message after the colon, but the person who sent the tweet read Ekramibrahim's message and retweeted it to her followers. This secondary message is the retweet, and the reader knows it was seen by at least the followers of Ekramibrahim and the person who retweeted it. A message can be retweeted an infinite number of times, though a user who sees a retweet only knows that at least one person retweeted it; in practice, most tweets are not retweeted, those that are are not retweeted often, and the retweet rate decays to almost 0% after 24 hours [Kwak et al., 2010, Liere, 2010, Starbird & Palen, 2012].[5]

A message can also contain a user mention or a link. A message that directly refers to another user by name is said to contain a user mention. If a user writes, "@ramezm i noticed a debate: #25jan or #jan25", @ramezm will receive a personal notification about the message; a tweet with a user mention is still viewable by the followers of the original author. A retweet is when a user shares the message of someone he follows with those who follow him; it is akin to forwarding an e-mail to everyone in your address book. Tweets also often contain links to photos and articles, though those messages are rarely retweeted [Suh et al., 2010].

Retweets, links, and mentions are not as effective at promoting coordination as hashtags. While more retweets of one tweet means that more people have seen the same set of information, the prevalence of hashtags means the information in a retweet is also available to those searching for hashtags that the retweet happen to contain. The same logic is true of links: if a link is meant to provide coordinating information, it will almost certainly contain a hashtag that is also relevant to coordination. While it is possible that user mentions have a strong coordinating effect outside of their employment of hashtags, they are dyadic and tend to be part of conversations - they are not used to mobilize protesters.

Equation 1.2 shows the calculation of these other measures of potential coor-

[5]With one extra click, a user can see how many times the original tweet was retweeted, but there is no way for the researcher to observe if a user knows how many times a tweet was retweeted.

dination. For each day t and country i, the measure counts the number of tweets, K, and tweets with measure M_k. m is a count of tweets with a link, mention of another user, or that are a retweet, depending on the measure in question. Dividing the measure by the number of tweets that day in that country quantifies the amount of other possible coordination that could have occurred in addition to $Coordination_{i,t}$.

$$MPercent_{i,t} = \sum_{i=1}^{16} \sum_{t=1}^{426} \frac{1}{K} * \sum_{k=1}^{K} m_k \tag{1.2}$$

Note that $Coordination_{i,t}$ is one variable that encompasses the two mechanisms, signalling and protest information provision. This measurement choice was made for 4 reasons. First, $Coordination_{i,t}$ should measure information protest information provision because it measures hashtag concentration, and individuals on Twitter use hashtags to quickly identify their tweets as being about a specific topic. During periods of heightened political awareness, the most common hashtags are most likely to be about politics; a tweet with "#jan25" is not likely to be about sports or the weather, for example. While this measure could create false positives - days that appear to have high coordination but are really people talking something else like a meme or a cultural event - the Supplementary Materials' case study shows that this is not the case.

Second, $Coordination_{i,t}$ is preferred to a topic model of protest information provision because it scales easily and is directly comparable across countries. Tweets often contain slang that varies by country, so making a topic model for each country is a large project in its own right. Aside from requiring much more labor, creating a topic model, whether unsupervised or not, risks constricting results to words or topics that the researcher has an *a priori* expectation will matter [Grimmer & Stewart, 2013]. A hashtag Gini, on the other hand, is agnostic to what words people say or how many topics they discuss; caring only about the hashtags, it will measure any hashtag used (not just the ones thought of in advance), revealing after the fact which hashtags are most salient.

Third, $Coordination_{i,t}$ is preferred to selecting specific hashtags because it scales easily and is directly comparable across countries. The hashtags used to coordinate an event are different in each country and change over time. Determining which hashtags to use for a given period of time therefore requires subject matter expertise on the digital arena of many countries over many months, which is not feasible for a large cross-national study. Moreover, because individuals pool on certain hashtags during protest periods, measuring hashtag concentration picks up on the most common protest hashtags. This behavior is explored in more detail in the Supplementary Materials.

Fourth, $Coordination_{i,t}$ is preferred over a direct measure of signalling is because it is not clear how to measure signalling on Twitter. he best approach would be to separate out the amount of signalling and protest information tweets which come from the periphery and the core. But it is not clear how to classify a “signal” tweet, and a more precise method of measuring protest information tweets does not scale well, as discussed in the previous paragraph. In addition, the three other coordinating behaviors discussed above can each signal peripheral participation. A link to a news article about the mass protests is not the same as a user saying, “I will protest tomorrow and I have never protested before”, even if both tweets come from the same account. Similarly with a photo showing a diverse crowd at a protest. In Table 1.4, an attempt is made to measure the signal component of coordination, and Section

1.5.2 discusses how to tease apart that $Coordination_{i,t}$ which is from the core.

1.5.2 Core Coordination

To measure coordination from the core, one has to first identify individuals at the center the network. Identifying this core is difficult. There are too many users - 20,094 in Bahrain and 79,235 in Egypt alone - in the Twitter data to assign manually an identity to each one, and that attempt would result in a low identification rate because Twitter does not require individuals to publicly disclose any identification information. One can measure, however, the number of followers each account has; this measure, in-degree centrality, is not as precise a measure of centrality as those created with complete network data [Kwak et al., 2010, Pei et al., 2014], but complete network data are not available. Those in the core are therefore approximated based on the distribution of popularity in each country. For the main model, a tweet belongs to a core member if its author's number of followers are at or above the 95^{th} percentile for all users in country i. More formally,

$$Core = \begin{cases} 1 & \text{if } \mathrm{PR}_i(\mathrm{f}) \geq .95 \\ 0 & \text{if } \mathrm{PR}_i(\mathrm{f}) < .95 \end{cases} \tag{1.3}$$

where PR(f) is the percentile ranking of the tweet based on the number of followers.

Previous work that manually identified a random sample of users from Tunisia and Egypt informed the selection of this threshold [Lotan et al., 2011]. Table 1.10 compares the number of followers and tweet production for the categories identified in Lotan et. al 2011 with this paper's primary popularity threshold; the threshold used later is bolded. In Tunisia, the core measure appears to roughly be most similar to bloggers; in Egypt, to bloggers and activists, though the manually identified accounts in Egypt are much more popular than any of the popularity measures. Notice though mainstream media account and employees of mainstream media are the most central in each country and skew the country-level results upwards. The Results section shows that varying the follower threshold does not change the result. The Supplementary Materials also show how tweet production and the ratio of the core's followers to the periphery's followers varies by country and threshold; users at the 95% threshold account for 10% of all tweets in Kuwait, up to 50% in Syria.

Table 1.1: Comparing Core Measure with Hand-coded Accounts*

Group	Accounts	Followers Avg.	Tweet Avg.	Mention %	Retweet %	Hashtag %	Link %
Egypt - Lotan	37	15138.71	949.78	0.40	0.06	0.36	0.32
Mainstream Media	1	103927.00	5281.00	0.00	0.00	0.74	0.70
Non-Media Org.	2	23877.40	457.50	0.32	0.07	0.22	0.52
MSM Employee	9	22463.50	650.22	0.41	0.01	0.21	0.21
Blogger	15	8394.17	1070.67	0.52	0.08	0.33	0.22
Activist	10	8036.55	703.40	0.42	0.07	0.28	0.33
Core 99.9 Percentile	80	37001.28	924.69	0.33	0.02	0.39	0.44
Core 99 Percentile	793	7104.31	736.38	0.45	0.05	0.27	0.32
Core 98 Percentile	1585	4033.08	591.71	0.46	0.05	0.25	0.31
Core 97 Percentile	2378	2875.48	515.21	0.46	0.04	0.25	0.31
Core 96 Percentile	3170	2256.03	453.52	0.45	0.04	0.25	0.30
Core 95 Percentile	**3962**	**1868.79**	**409.94**	**0.45**	**0.04**	**0.25**	**0.30**
Blackout	**740**	**8046.33**	**650.95**	**0.22**	**0.05**	**0.24**	**0.61**
Tunisia - Lotan	10	7942.94	248.60	0.33	0.12	0.59	0.56
Mainstream Media	2	5604.49	741.00	0.16	0.12	0.77	0.78
MSM Employee	1	52503.00	1.00	0.00	0.00	0.00	1.00
Blogger	3	1910.77	258.33	0.57	0.13	0.30	0.20
Activist	4	2496.28	57.00	0.59	0.09	0.36	0.29
Core 99.9 Percentile	7	17749.31	206.71	0.25	0.06	0.22	0.68
Core 99 Percentile	62	4880.14	410.92	0.37	0.06	0.27	0.55
Core 98 Percentile	123	3095.44	444.93	0.44	0.14	0.31	0.43
Core 97 Percentile	184	2392.96	374.43	0.45	0.13	0.31	0.40
Core 96 Percentile	245	1968.22	337.18	0.47	0.12	0.30	0.37
Core 95 Percentile	**307**	**1681.87**	**308.22**	**0.47**	**0.11**	**0.31**	**0.38**

*Categories are borrowed from Lotan et. al 2011. They coded for accounts associated with mainstream media organizations, mainstream new media organizations (news sites that exist only online) mainstream media employees, any organization that is not a media organization (Vodafone and Wikileaks are their examples), bloggers, activists, digerati, political actors, celebrities, researchers, bots, and a residual category. Any of those categories not identified here means that no account from that category was found in the data.

*The bold rows represent the category used to identify core members. Other categories are used in robustness checks, with no changes to the results. For a discussion of the Blackout row, please see Section 1.8.

Having identified tweets produced from those in the core, one can then identify when the core engages in coordination. Because hashtags are the primary method of coordination and high levels of coordination lead to protest, the percentage of hashtags per country per day produced from the core is interacted that with the coordination measure. The percent of tweets with hashtags that are created in the core is defined as:

$$Core\ Coordination_{i,t} = \sum_{i=1}^{16}\sum_{t=1}^{426} Coordination_{i,t} * \frac{1}{K} * \sum_{k=1}^{K} Core_k * Hashtag_k \quad (1.4)$$

For each country i on each day t, each of the K tweets is read for whether it contains a hashtag and is from a core account. The number of those tweets is divided by the number of tweets in that country-day and interacted that with that country-day's level of coordination, resulting in a core coordination measure for that country-day. The regression results leave the constituent parts of the variable as the variable name to ease interpretation; the summary statistics use the shortened name to save space.

In-degree - the number of followers an account has - is chosen to measure core position for 3 reasons. First, it is the only network centrality measure which does not require complete network data. Twitter imposes limits on how often one can download data, making it impossible to create the adjacency matrix necessary to create measures like Eigenvector or k-core centrality. Second, one could approximate a complete network by inferring ties when a retweet or mention occurs, but that requires pooling data, losing the time component necessary for this theory [Barberá, Wang, Bonneau, Jost, Nagler, Tucker & González-Bailón, 2015]. Finding "hidden influentials", those who follow many more people than follow them (normal Twitter behavior) and who are mentioned much more than they mention others (happens for well-known accounts like celebrities or politicians) is another strategy, but it again relies on pooling across time [Gonzalez-Bailon, Borge-Holthoefer & Moreno, 2013]. Third, this paper's theory is about how position in a social network correlates with protest mobilization. If in-degree does not identify those who influence protest, then that means those in the network core do not influence protest. Future work should then find ways to identify individual accounts that are influential; one approach, handcoding accounts by profession, is explored in Table 1.5.

1.6 Model

The base model is:

$$Protests_{i,t} = \beta_0 + \beta_1 * \Omega_{i,t-1} + \boldsymbol{\beta} * \mathbf{X}_{i,t-1} + Protests_{i,t-1} + \epsilon_{i,t} \tag{1.5}$$

where Ω represents the independent variables of interest in each model, $\mathbf{X}$ represents a series of controls, and ϵ is a stochastic error term. Because the dependent variable is a count of protests, it is an integer always greater than or equal to 0. Since $Protests_{i,t}$ is overdispersed and the zeroes are true zeroes, a negative binomial model instead of a Poisson or zero-inflated negative binomial is used.

Because high levels of coordination are colinear with high levels of hashtag usage, the model controls for the percent of a day's tweets that have hashtags, ensuring that it measures actual coordination and not a coincidental increase in hashtag usage. The models of peripheral coordination control for the percent of a day's tweets which are retweets, contain links, or mention another user because those features may have some coordination effect. The models of core coordination similarly control for the percent of all tweets with at least one hashtag that are from accounts of the core; the percent of all tweets that are retweets which are from core members; and so on for links and mentions.

There are three non-individual controls: country fixed-effects, a lagged dependent variable, and a lagged measure of the number of repression events as measured by ICEWS. Repression is any event with a CAMEO of exhibiting military posture (event root code 15), coercion (17), using unspecified unconventional violence (18), a physical assault (182), torture (1822), or death by physical assault (1823).

Every variable on the right-hand side is lagged by one day to mitigate any simultaneity effects. All models include country fixed effects but no day fixed effects, as the latter bias the errors and lead to underestimates of protests. Finally, all models are run with country-clustered standard errors.

Table 1.2 shows the correlation between the main independent variables, and Table 1.3 shows the average value of each variable per country (along with each's total tweets and protests).

Table 1.2: Variable Correlation

	Protest$_{i,t}$	Coord.$_{i,t-1}$	Hashtag %$_{i,t-1}$	Retweet %$_{i,t-1}$	Link %$_{i,t-1}$	Mention %$_{i,t-1}$	Protest$_{i,t-1}$	Core Hashtag%$_{i,t-1}$	Core Retweet%$_{i,t-1}$	Core Link%$_{i,t-1}$	Core Mention%$_{i,t-1}$	Repression$_{i,t}$	Core Coord.$_{i,t-1}$
Protest$_{i,t}$	1	0.281	0.218	-0.012	0.111	-0.155	0.785	0.178	0.099	0.149	0.1	0.571	0.309
Coord.$_{i,t-1}$	-	1	0.594	-0.038	0.101	-0.237	0.295	0.535	0.372	0.519	0.349	0.269	0.88
Hashtag %$_{i,t-1}$	-	-	1	0.203	0.427	-0.448	0.231	0.521	0.351	0.405	0.186	0.185	0.649
Retweet %$_{i,t-1}$	-	-	-	1	0.11	-0.233	-0.013	0.122	0.279	0.156	0.098	0.029	-0.051
Link %$_{i,t-1}$	-	-	-	-	1	-0.659	0.106	0.325	0.232	0.23	0.024	0.063	0.291
Mention %$_{i,t-1}$	-	-	-	-	-	1	-0.155	-0.344	-0.265	-0.3	0	-0.126	-0.351
Protest$_{i,t-1}$	-	-	-	-	-	-	1	0.175	0.108	0.152	0.1	0.673	0.317
Core Hashtag %$_{i,t-1}$	-	-	-	-	-	-	-	1	0.656	0.702	0.545	0.158	0.729
Core Retweet %$_{i,t-1}$	-	-	-	-	-	-	-	-	1	0.583	0.501	0.152	0.467
Core Link %$_{i,t-1}$	-	-	-	-	-	-	-	-	-	1	0.49	0.142	0.615
Core Mention %$_{i,t-1}$	-	-	-	-	-	-	-	-	-	-	1	0.155	0.417
Repression$_{i,t-1}$	-	-	-	-	-	-	-	-	-	-	-	1	0.271
Core Coord.$_{i,t-1}$	-	-	-	-	-	-	-	-	-	-	-		1

Table 1.3: Variables by Country

Country$_i$	Protests$_{i,t}$	Tweets$_{i,t}$	Coord.$_{i,t}$	Hashtag %$_{i,t}$	Retweet %$_{i,t}$	Link %$_{i,t}$	Mention %$_{i,t}$	Core Hashtag %$_{i,t}$	Core Retweet %$_{i,t}$	Core Link %$_{i,t}$	Core Mention %$_{i,t}$	Core Coord.$_{i,t-1}$
Egypt	3,379	3,742,648	0.59	0.21	0.04	0.29	0.42	0.52	0.55	0.47	0.48	0.31
Syria	2,057	229,476	0.60	0.34	0.03	0.61	0.22	0.54	0.49	0.48	0.35	0.37
Yemen	1,885	61,517	0.39	0.25	0.05	0.62	0.19	0.52	0.23	0.41	0.24	0.25
Tunisia	882	228,554	0.37	0.25	0.07	0.43	0.40	0.50	0.61	0.37	0.49	0.19
Bahrain	798	1,056,990	0.53	0.22	0.06	0.17	0.42	0.39	0.28	0.46	0.28	0.22
Libya	663	84,991	0.37	0.24	0.11	0.34	0.32	0.42	0.54	0.37	0.38	0.17
Iraq	585	146,113	0.35	0.18	0.09	0.28	0.35	0.44	0.51	0.31	0.43	0.16
Jordan	511	273,227	0.38	0.21	0.06	0.45	0.36	0.45	0.48	0.42	0.38	0.17
Morocco	298	300,454	0.25	0.18	0.07	0.34	0.43	0.40	0.40	0.33	0.34	0.11
Lebanon	261	522,891	0.36	0.21	0.06	0.28	0.41	0.50	0.61	0.50	0.41	0.18
Algeria	248	7,474	0.08	0.21	0.11	0.44	0.34	0.40	0.43	0.36	0.18	0.04
Kuwait	161	29,838	0.09	0.09	0.02	0.13	0.52	0.09	0.02	0.06	0.21	0.01
Saudi Arabia	156	4,425,797	0.48	0.13	0.05	0.16	0.51	0.46	0.37	0.45	0.44	0.23
Oman	150	8,509	0.02	0.09	0.05	0.31	0.48	0.06	0.02	0.11	0.14	0.00
UAE	58	1,531,524	0.35	0.15	0.07	0.20	0.43	0.36	0.37	0.42	0.31	0.13
Qatar	29	1,104,995	0.34	0.12	0.06	0.10	0.43	0.39	0.35	0.37	0.32	0.14

1.7 Results

The main results are presented in Table 1.4. Columns 1-2 show coordination only from the periphery, and Columns 3-4 build the models for core coordination in the same way. The main model, used throughout the rest of the paper, is shown in Column 5. Across all models, $Coordination_{i,t-1}$, the measure for peripheral coordination, is significant with a p-value much less than .01. The only only other significant variables are a lagged dependent variable (positive), intercept (negative), lagged repression (weakly positive), and non-influential hashtag percent (positive). Note that Model 5, the full model, suggests that coordination from the core is inversely associated with protests.

Table 1.4: Peripheral Coordination and Protest

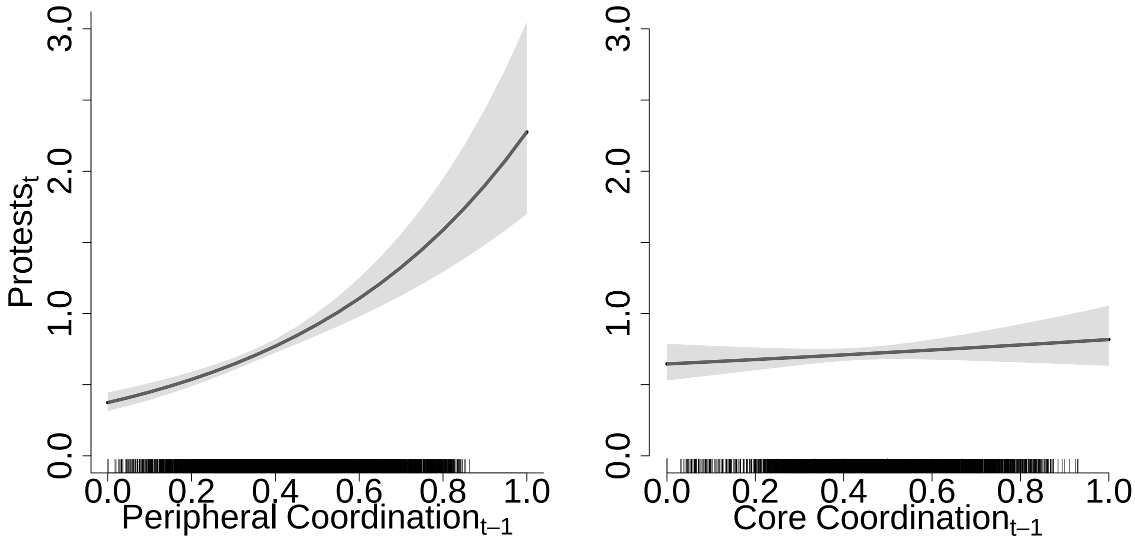
	DV: $Protest_{i,t}$				
	Coordination		Core Coordination		Full Model
	(1)	(2)	(3)	(4)	(5)
$\text{Coordination}_{i,t-1}$	1.932***	1.936***	1.809**	1.830**	2.575***
	(0.472)	(0.469)	(0.766)	(0.774)	(0.639)
$\text{Hashtag }\%_{i,t-1}$		0.707			0.578
		(0.539)			(0.616)
$\text{Retweet }\%_{i,t-1}$		0.405			−0.362
		(0.876)			(1.019)
$\text{Link }\%_{i,t-1}$		−0.536			−0.763**
		(0.386)			(0.378)
$\text{Mention }\%_{i,t-1}$		−0.858*			−0.921**
		(0.512)			(0.410)
$\text{Repression}_{i,t-1}$	0.020*	0.020*	0.021*	0.022**	0.021*
	(0.011)	(0.012)	(0.011)	(0.011)	(0.012)
$\text{Protest}_{i,t-1}$	0.127***	0.121***	0.126***	0.125***	0.116***
	(0.010)	(0.011)	(0.010)	(0.010)	(0.010)
$\text{Core Hashtag }\%_{i,t-1}$			0.600	0.631	0.900**
			(0.393)	(0.441)	(0.372)
$\text{Core Retweet }\%_{i,t-1}$				0.063	0.158
				(0.398)	(0.413)
$\text{Core Link }\%_{i,t-1}$				0.450	0.711
				(0.553)	(0.565)
$\text{Core Mention }\%_{i,t-1}$				−0.485	−0.159
				(0.316)	(0.258)
$\text{Coordination}_{i,t-1}\text{*Core Hashtag }\%_{i,t-1}$			−0.313	−0.503	−1.868**
			(1.302)	(1.333)	(0.942)
Intercept	−0.934***	−0.569**	−1.182***	−1.295***	−1.019***
	(0.044)	(0.229)	(0.160)	(0.215)	(0.207)
Country FE	Yes	Yes	Yes	Yes	Yes
N	6,800	6,620	6,800	6,800	6,620
Log Likelihood	−8,469.486	−8,296.844	−8,463.422	−8,459.476	−8,280.817

*p < .1; **p < .05; ***p < .01
Model 3 includes $Coordination_{i,t}$ and $Core\ Hashtag_{i,t-1}$ because those are the constituent parts of the measure of influential coordination.

Because the model is not linear, coefficients do not directly translate into changes in the outcome variable. The marginal effects of the two variables are shown in Figure 1.2. Going from no coordination to the maximum observed values leads to about 2 additional protests, a 400% increase, while there exists no effect for core coordination.

A series of time series diagnostic tests confirm the model specification.[6] A

[6]A panel OLS model is used for the diagnostics to ease calculation of the test statistics. The robustness checks that the panel OLS results match those of the negative binomial.



(a) Peripheral Coordination (b) Core Coordination

Figure 1.2: Marginal Effects of Peripheral and Core Coordination

Durbin-Watson test for serial correlation returns a test statistic of 1.9741 and p-value of .1303, suggesting no serial correlation. The Dickey-Fuller coefficient is -11.25 and has a p-value less than .01, so the dependent variable is stationary (visual inspection also confirms the stationarity). The Breusch-Pagan test statistic is 539.09 with a p-value almost at 0; to control for the heteroskedasticity, I use country-clustered standard errors. Finally, a Lagrange-Multiplier test with the King & Wu test for two-way fixed effects returns a chi-square value of -.81, so it is safe to avoid using time fixed effects.

1.7.1 Verification

There are 3 possibilities that may undercut these findings. First, the models may use the wrong measure of coordination, both for the core and the periphery. Second, the operationalization of core members may be wrong. Third, reliance on machine-coded data may bias in favor of finding results.

Figure 1.3 allays the first concern. To confirm that $Coordination_{i,t}$ measures coordination, a topic model for Egypt and Bahrain was created. Those two countries were chosen because they experienced widespread protest and have too many tweets to code individually. For each country, 3,000 tweets were randomly selected and coded into overlapping categories, one of which was protest information. A naïve bayes classifier was trained on a random 95% of each country's coded tweets; the other 5% was used to validate out of sample performance. This process was repeated 30 times, and the results were averaged into a final model; this process is known as bagging and is akin to bootstrapping in regression. The resulting model is applied to each country's tweets, creating a classification for every single tweet in the sample from Egypt and Bahrain. Once protest information tweets are identified, they are aggregated by country-day and compared to the $Coordination_{i,t}$ measure. That result is shown in Figure 1.3: there is a strong positive relationship between the measure of coordination and the actual number of protest information tweets. See the Supplementary Materials for the codebook and more details on the topic model.

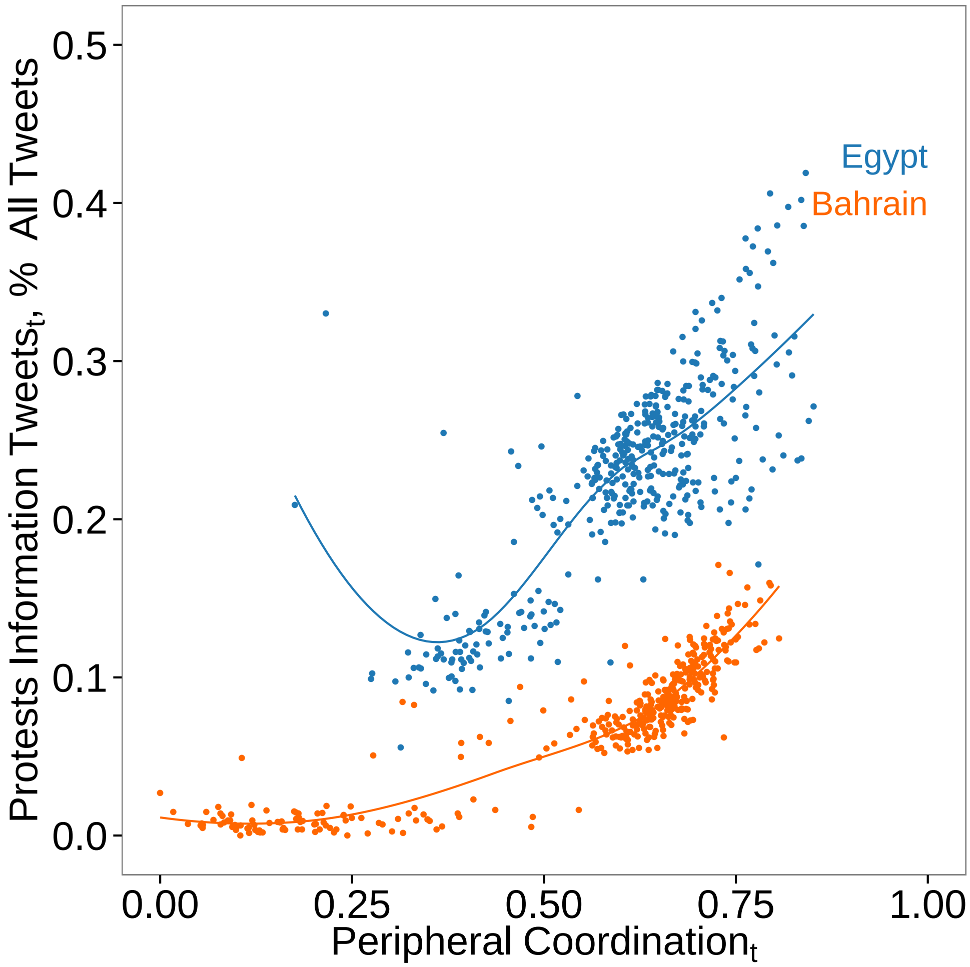


Figure 1.3: Verifying Operationalization of Coordination

Two placebo tests also address concerns about $Coordination_{i,t}$. The first test shows that the correlation of $Coordination_{i,t-1}$ and $Protest_{i,t}$ decreases substantially as an increasing number of the most popular hashtags on a country-day are removed, suggesting that coordination on non-protest hashtags does not drive subsequent protest. The second test shows that the correlation of $Coordination_{i,t}$ and $Protest_{i,t}$ peaks with a 1-day lag and decreases monotonically before and after. The Supplementary Materials present the tests in more detail.

Table 1.5 addresses the second concern. In column 1, accounts from Lotan et. al 2011 in this sample are identified and controlled for. Accounts from activists or bloggers are called "Online Actors", while there are not variables for politician, researcher, digerati, or celebrity accounts because those were not found in the sample from either country. It appears that mainstream media accounts (official accounts of news organizations) do positively correlate with subsequent protest. The main coordination measure is still strongly significant, as is retweet percentage; links appear to decrease in rate leading up to protests. The second and third columns of Table 1.5 show alternative measures of coordination in the core. *Core Reachout* $\%_{i,t-1}$ measures the percent of all retweets and mentions that come from those in the core. Core coordination may occur through those in the core engaging with specific individuals (mentions) or acting as information brokers (finding important tweets and retweeting them), not through hashtags. Column 2 controls for this possibility. On the other hand, the core may have a coordination effect simply by being active leading up to protests; their activity may signal a breakdown of support for the regime, a willingness to incur high personal costs that inspires the periphery, or they just may not use hashtags. There appears to be some effect for tweets from the core - *Core Tweet* $\%_{i,t-1}$ is significant at $p \leq 1$. - but their tweet activity on high coordination days does not correlate with subsequent protest. In all three models, coordination from the periphery is still significant and coordination from the core is not.

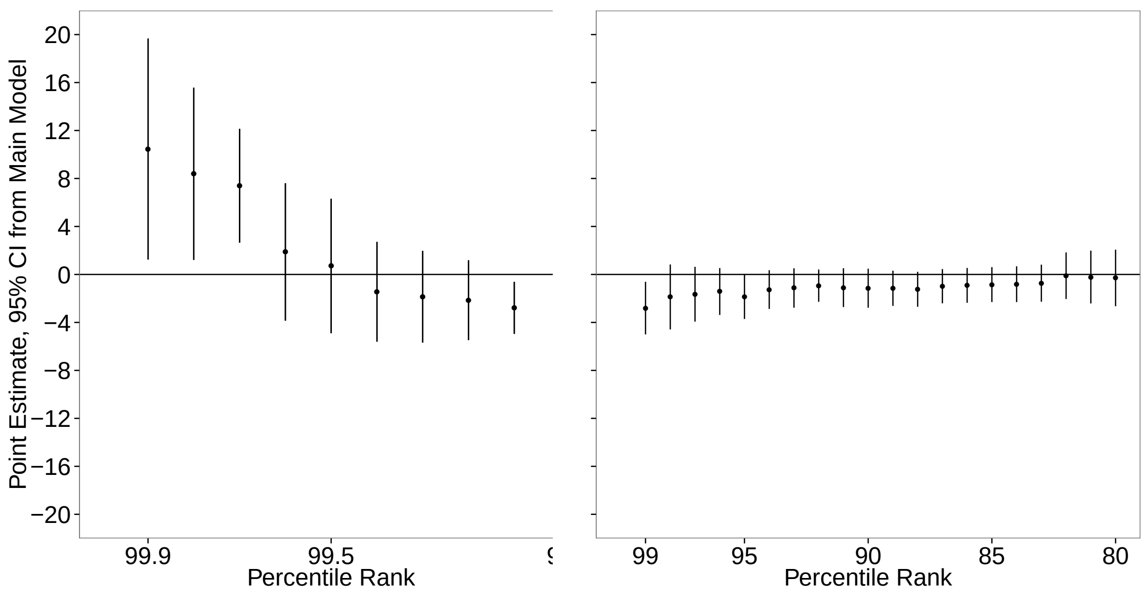
Model 4 of Table 1.5 tests Hypotheses 1a and 1b. Whereas the rest of the models presented in this paper can only test coordination overall, Hypothesis 1, here an attempt is made to separate out the effects of the size of the periphery (H1a) and its contribution to coordination (H1b). To test H1a, I measure the number of unique accounts per day from the periphery and divide it by the total number of accounts that day. The resulting percentage operationalizes the idea that a critical mass or a social network is required to start collective action [Marwell, Oliver & Prahl, 1988]. Model 4 of Table 1.5 suggests this is not the case, and its inclusion does not change the results on other variables.

Table 1.5: Robust to Operationalization of Core, Periphery

	$Protest_{i,t}$			
	Core Manual ID	Core Reachout	Core Tweet Share	Critical Mass
	(1)	(2)	(3)	(4)
$\text{Coordination}_{i,t-1}$	3.041***	1.678**	2.308***	2.920***
	(0.419)	(0.729)	(0.772)	(0.658)
$\text{Hashtag } \%_{i,t-1}$	−1.728	0.412	0.448	0.509
	(7.574)	(0.641)	(0.620)	(0.638)
$\text{Retweet } \%_{i,t-1}$	12.705***	−0.130	−0.246	−0.439
	(1.407)	(0.926)	(0.934)	(1.007)
$\text{Link } \%_{i,t-1}$	−1.525	−0.816**	−0.852**	−0.884**
	(1.341)	(0.359)	(0.374)	(0.400)
$\text{Mention } \%_{i,t-1}$	−1.596	−0.701	−0.941**	−0.876**
	(6.119)	(0.460)	(0.433)	(0.408)
$\text{Mainstream Media } \%_{i,t-1}$	18.313***			
	(5.636)			
$\text{MSM Empl. } \%_{i,t-1}$	62.784			
	(53.143)			
$\text{Online Actor } \%_{i,t-1}$	−7.310**			
	(3.191)			
$\text{Spam } \%_{i,t-1}$	−7.989			
	(17.187)			
$\text{Critical Mass}\%_{i,t-1}$				0.704
				(1.963)
$\text{Repression}_{i,t-1}$	0.016***	0.020*	0.021*	0.022*
	(0.005)	(0.011)	(0.011)	(0.012)
$\text{Protest}_{i,t-1}$	0.044***	0.118***	0.116***	0.115***
	(0.009)	(0.010)	(0.010)	(0.010)
$\text{Core Reachout } \%_{i,t-1}$		1.358		
		(1.247)		
$\text{Core Tweet } \%_{i,t-1}$			1.078*	
			(0.608)	
$\text{Core Hashtag } \%_{i,t-1}$		0.416	0.366	0.785**
		(0.514)	(0.518)	(0.321)
$\text{Core Retweet } \%_{i,t-1}$		−0.233	0.101	0.179
		(0.580)	(0.419)	(0.392)
$\text{Core Link } \%_{i,t-1}$		0.512	0.365	0.569
		(0.528)	(0.650)	(0.695)
$\text{Core Mention } \%_{i,t-1}$		−1.394	−0.529	−0.478*
		(1.092)	(0.370)	(0.290)
$\text{Coordination}_{i,t-1}\text{*Core Reachout } \%_{i,t-1}$		0.595		
		(1.074)		
$\text{Coordination}_{i,t-1}\text{*Core Tweet } \%_{i,t-1}$			−1.276	
			(1.246)	
Intercept	0.335	−0.848***	−0.832***	−1.760
	(4.195)	(0.212)	(0.162)	(2.026)
Country FE	Yes	Yes	Yes	Yes
N	830	6,620	6,620	6,620
Log Likelihood	−1,712.510	−8,282.427	−8,282.744	−8,275.298

*p < .1; **p < .05; ***p < .01

To confirm the 95% threshold used to identity the core, the threshold was varied from the 80^{th} percentile to the 99.9^{th}, and Model 5 from Table 1.4 is rerun for each threshold. Figure 1.4 shows how the significance level of $Coordination_{i,t-1}$ *$Core\ Hashtag\ \%_{i,t-1}$ varies as the percentile threshold changes; the horizontal lines are at ± 1.96 to show significance at the 5% level. Figure 1.4a shows the result from 10 models, where model 1 uses a cutoff at the 99.9^{th} percentile and model 10 is at the 99^{th}; Figure 1.4b shows the result from 20 models, where model 1 uses a 99^{th} percentile threshold and the 20^{th} uses the 80^{th} percentile. In all iterations, $Coordination_{i,t-1}$, $Repression_{i,t-1}$, and $Protest_{i,t-1}$ remain significant.



(a) Top 1%

(b) 99th - 80th Percentile

Figure 1.4: Change in Effect Size as Function of Core Threshold

The results in Figure 1.4 are intriguing. The effect of core coordination hovers around 0 for most of threshold's range and is distinguishable from zero at the 99^{th} and 99.1^{st} percentiles, as well as at the 95^{th} percentile. On the other hand, at the upper extreme of the follower distribution, the 99.7^{th} percentile and above, the sign on $Core\ Coordination_{i,t-1}$ is positive and significant using a 95% confidence interval.

The apparent positive effect from core coordination above the 99.7^{th} percentile should not, however, be assigned much weight, for 3 reasons. First, these models also find that *Core Link* $\%_{i,t-1}$ is negative and significant, with an effect from half as strong as to equal with that of core coordination. Second, the pooled results are driven by outlier countries with few tweets and users at or above the 99.7^{th} percentile threshold. Whether the threshold is 99.7, 99.8, or 99.9, the effect disappears when the model is rerun using only Bahrain, Egypt, Jordan, Lebanon, Qatar, Saudi Arabia, and the United Arab Emirates, the only countries with more than 10,000 tweets from users at or above the 99.7^{th} percentile. All these countries are also the only ones with more than 25 users at this point in the distribution (except for Morocco, with 30). Because the final dataset is at the country-day level, it does not distinguish between a day in Egypt that may have 500 tweets from 20 accounts in the 99.7^{th} percentile core from one in Algeria that has 1 tweets from 1 of the 2 accounts above the same cut-off. The resulting models therefore overweight Algerian core users. Rerunning the main model with only the 7 countries just described therefore provides a more accurate understanding of the dynamics this far into the followers' distribution, and models on these 7 countries show no effect for members of the core. Third, it is more likely than not that these accounts represent institutions such as news outlets or non-profit organizations than people.[7] As Table 1.6 shows, these accounts are frequent tweeters, and those tweets are more likely to contain hashtags or links. Such behavior is most similar to how news organizations use Twitter [Lotan et al., 2011]; Table 1.10 in Section1.5.1 compares the 99.9^{th} percentile core users to confirmed news accounts in Tunisia and Egypt, showing similarity between the two. The probability

[7] I cannot know for sure because the data are anonymized.

that these accounts are news organizations is further increased by rerunning the main model using only Arabic tweets. Using only Arabic tweets when the core is defined at the 99.7^{th}, 99.8^{th}, or 99.9^{th} thresholds, $Coordination_{i,t-1}$ *$Core\ Hashtag\ \%_{i,t-1}$ is not statistically significant. Overall, the positive effect suggested in Figure 1.4a is probably driven by a few media accounts in countries with less Twitter data than others in the sample.

Table 1.6: Core Threshold Descriptive Statistics Across Countries

Country	Group	Accounts	Followers Avg.	Tweet Avg.	Mention %	Retweet %	Hashtag %	Link %
Algeria	99.7^{th} Percentile	2	6786.50	1.50	0.33	0.00	0.33	1.00
Bahrain	99.7^{th} Percentile	61	15294.42	530.87	0.17	0.03	0.57	0.49
Egypt	99.7^{th} Percentile	238	17975.93	992.12	0.37	0.06	0.32	0.42
Iraq	99.7^{th} Percentile	14	23258.58	455.00	0.74	0.07	0.29	0.05
Jordan	99.7^{th} Percentile	26	9197.23	436.38	0.19	0.03	0.35	0.73
Kuwait	99.7^{th} Percentile	10	25240.53	7.00	0.40	0.07	0.07	0.30
Lebanon	99.7^{th} Percentile	52	7997.13	550.71	0.19	0.15	0.26	0.63
Libya	99.7^{th} Percentile	11	17287.99	276.36	0.39	0.05	0.59	0.25
Morocco	99.7^{th} Percentile	30	19132.81	165.03	0.45	0.03	0.27	0.37
Oman	99.7^{th} Percentile	2	122889.50	1.50	0.67	0.00	0.00	0.33
Qatar	99.7^{th} Percentile	65	23668.00	428.31	0.42	0.06	0.29	0.33
Saudi Arabia	99.7^{th} Percentile	266	13375.55	588.77	0.61	0.03	0.19	0.24
Syria	99.7^{th} Percentile	11	5484.73	630.27	0.14	0.02	0.39	0.74
Tunisia	99.7^{th} Percentile	19	10068.36	377.95	0.34	0.04	0.31	0.63
UAE	99.7^{th} Percentile	142	22961.05	261.39	0.33	0.05	0.32	0.36
Yemen	99.7^{th} Percentile	5	5132.13	371.60	0.05	0.00	0.33	0.75

Table 1.7 verifies the ICEWS dependent variable. ICEWS relies on news reports, and these reports have well-known biases in coverage [Davenport & Ball, 2002, Herkenrath & Knoll, 2011, Eck, 2012]. Machine-coded events data can suffer from event duplication [Caren, 2014, Hammond & Weidmann, 2014], though one of ICEWS' strengths is its focus on event deduplication. The results could therefore be driven by news media's bias towards major, unexpected events and duplicated events. In the first column of Table 1.7, ICEWS' count of public statements is the dependent variable. If ICEWS simply picks up news activity, it should record a surge in public statements along with protest, and coordination will then correlate with public statements. Column 1 shows that coordination does not correlate with public statements, suggesting that $Protest_{i,t}$ actually captures protest.[8]

[8] Columns 1 and 2 use an ordinary-least squares estimator because the dependent variable is no longer a count.

The second column shows that the results do not appear to be driven by duplication. The dependent variable is *Protest Rate*$_{i,t}$, the number of protests on a country-day divided by the number of ICEWS events at that time. If ICEWS duplicates, then the protest rate should not change across the sample and there will be no correlation between *Coordination*$_{i,t-1}$ and *Protest Rate*$_{i,t}$. *Coordination*$_{i,t-1}$ is still significant and *Coordination*$_{i,t-1}$***Core Hashtag %*$_{i,t-1}$, in line with the main results.[9]

[9] The Supplementary Materials visualizes the protest rate and shows that it varies in tandem with real-world events.

Table 1.7: Robust to Dependent Variable

	Public Statements Rate$_{i,t}$	*Protest Rate*$_{i,t}$	*Protest*$_{i,t}$	
	All	All	Drop 3 SD	Drop Top Quarter
	(1)	(2)	(3)	(4)
Coordination$_{i,t-1}$	0.047	0.100***	2.052***	1.498*
	(0.039)	(0.036)	(0.567)	(0.794)
Hashtag %$_{i,t-1}$	−0.051***	0.058	1.127*	0.408
	(0.013)	(0.026)	(0.582)	(0.579)
Retweet %$_{i,t-1}$	0.002	−0.035	−0.628	−0.755
	(0.032)	(0.035)	(0.749)	(0.587)
Link %$_{i,t-1}$	0.029**	−0.035	−0.069	0.671
	(0.013)	(0.025)	(0.385)	(0.464)
Mention %$_{i,t-1}$	0.003	0.003	0.242	0.726
	(0.021)	(0.021)	(0.600)	(0.561)
Public Statements Rate$_{i,t-1}$ 0.099***				
	(0.013)			
Repression Rate$_{i,t-1}$		0.047**		
		(0.020)		
Repression$_{i,t-1}$			0.036***	0.056***
			(0.016)	(0.020)
Protest Rate$_{i,t-1}$		0.356***		
		(0.035)		
Protest$_{i,t-1}$			0.122***	0.063***
			(0.011)	(0.010)
Core Hashtag %$_{i,t-1}$	−0.00002	0.029***	0.301	−0.625
	(0.031)	(0.018)	(0.271)	(0.697)
Core Retweet %$_{i,t-1}$	0.011	−0.001	0.112	−0.122
	(0.007)	(0.014)	(0.295)	(0.117)
Core Link %$_{i,t-1}$	0.036	−0.001	0.972**	1.252**
	(0.023)	(0.014)	(0.405)	(0.622)
Core Mention %$_{i,t-1}$	−0.004	−0.006	−0.227	−0.136
	(0.025)	(0.017)	(0.282)	(0.333)
Coordination$_{i,t-1}$*Core Hashtag %$_{i,t-1}$	−0.021	−0.068	−1.291	1.119***
	(0.059)	(0.064)	(0.973)	(0.343)
Intercept	0.069***	0.041**	−2.062***	−3.783***
	(0.021)	(0.016)	(0.368)	(0.497)
Model	OLS	OLS	Neg. Binom.	Neg. Binom.
Country FE	Yes	Yes	Yes	Yes
N	6,620	6,620	6,471	2,916
R^2	0.091	0.221		
Adjusted R^2	0.088	0.218		
Log Likelihood			−7,244.759	−1,765.218

*p < .1; **p < .05; ***p < .01

Columns 3 and 4 in Table 1.7 present the final verification of ICEWS' protest count. Since newspapers overreport major events and ICEWS over reports newspapers, it is possible that the results are driven by the upper end of the protest distribution. Column 3 drops all protest-days that have protests 3 standard deviations above the country's average, and Column 4 drops all protest-days in the upper quartile of each country's protest distribution. In both cases, $Coordination_{i,t-1}$ is significant while elite coordination is not. The main finding of this paper, that coordination occurs along the periphery of a network, is not an artifact of using machine-coded data.[10]

In results shown in the Supplementary Materials, the main model is run while throwing out countries that may have overly influenced results. Removing the 5 countries with the highest levels of protests per capita, the results hold. Removing the 5 countries with the lowest levels of protests per capita, the results hold. Removing the 5 countries with the most tweets per capita, the results hold. Removing the 5 countries with the fewest tweets per capita, the results hold. These results suggest that the coordination patterns are widespread throughout the dataset and not dependent on a few countries.

Further verification of the dependent variable is shown in Figure 1.5, which shows that ICEWS' count of protests strongly correlates with hand-coded data. The

[10]Two new machine-coded projects, the University of Illinois' Social, Political, Economic Event Database (SPEED) and the Open Event Data Alliance's Phoenix project, look to improve on are exciting events data projects. SPEED combines machine-coded data with human verification to achieve human-level accuracy with machine-coded breadth [Nardulli, Althaus & Hayes, N.d.]. The Phoenix project, an open source system associated with Pennsylvania State University and Parus Analytical Systems, is a major evolution of the TABARI system. Phoenix's main advantage over ICEWS, which also uses a heavily modified version of TABARI, is that it releases new data daily, while ICEWS releases monthly on a one year delay. As of this writing, SPEED's public data only go through 2005, and Phoenix's data starts on June 20^{th}, 2014. Phoenix's Github page is https://github.com/openeventdata.

Armed Conflict Location and Event Dataset (ACLED) is hand-coded contains data the number of riots and protests in Algeria, Egypt, Libya, Morocco, and Tunisia for 2010 and 2011 [Raleigh et al., 2010]. ACLED provides greater event granularity than the Social Conflict in Africa Dataset, another hand-coded events dataset that contains protests [Hendrix et al., 2012]. The two measures have a Pearson's correlation coefficient of .468.

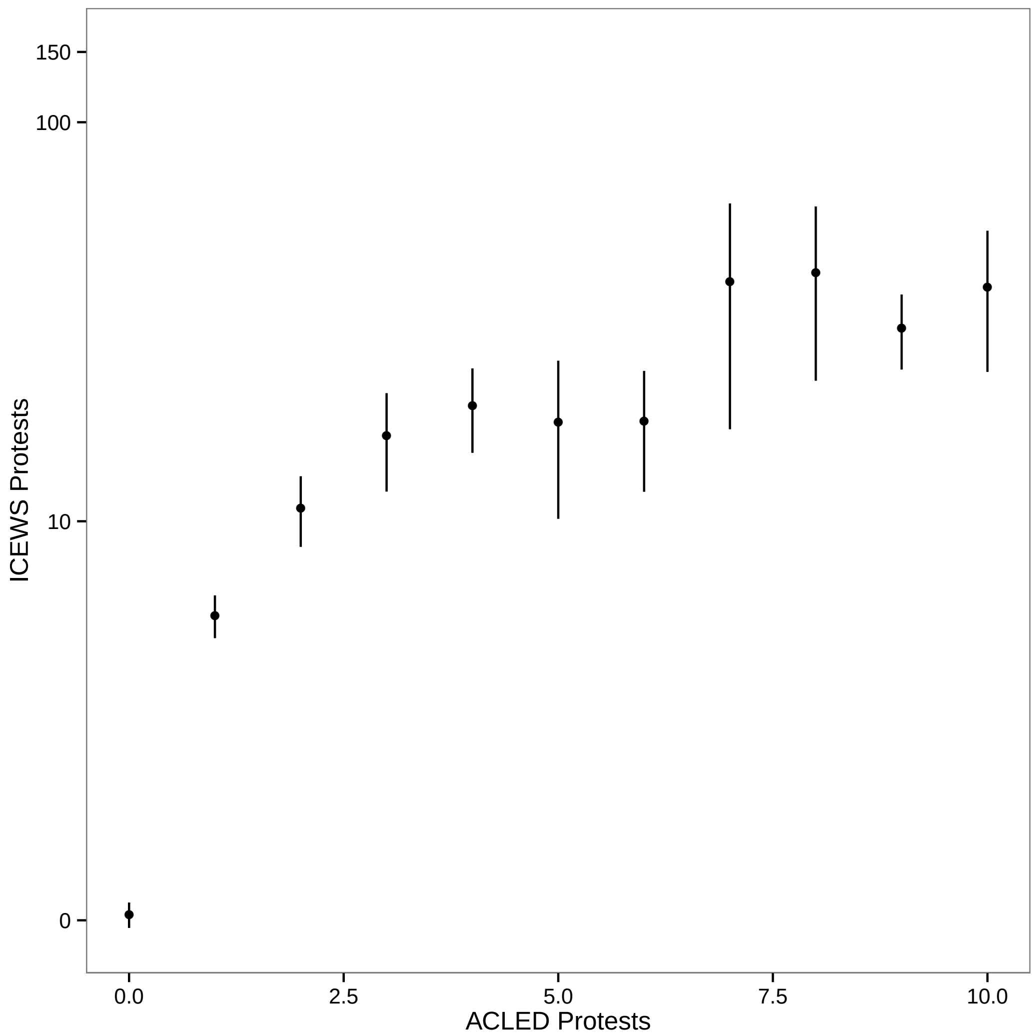


Figure 1.5: ICEWS Correlates with Handcoded Data

The Supplementary Materials use the ACLED measures as a dependent variable and shows that coordination may still occur without the core's coordination, though ACLED's little variation on the dependent variable means most results do not attain traditional levels of statistical significance. In results not presented, ICEWS is shown to correlate with GDELT [Leetaru & Schrodt, 2013], another machine-coded events dataset; their Pearson correlation coefficient is .785. All models presented here and in the Supplementary Materials were rerun using GDELT, and all results hold.

In the Supplementary Materials, I address other concerns, such as outliers and language use, about the data generating process; all results hold.

1.8 Exogenous Identification of the Core in Egypt

This section takes advantage of a sudden increase in the difficulty of accessing the internet in Egypt to identify core members of the Egyptian social network. Those who could access the internet during this period are more likely to be in the core than those who could not, so tracking their communicative behavior throughout the study should more precisely identify any role for the core. Those in the core according to this identification strategy also do not lead to protest mobilization.

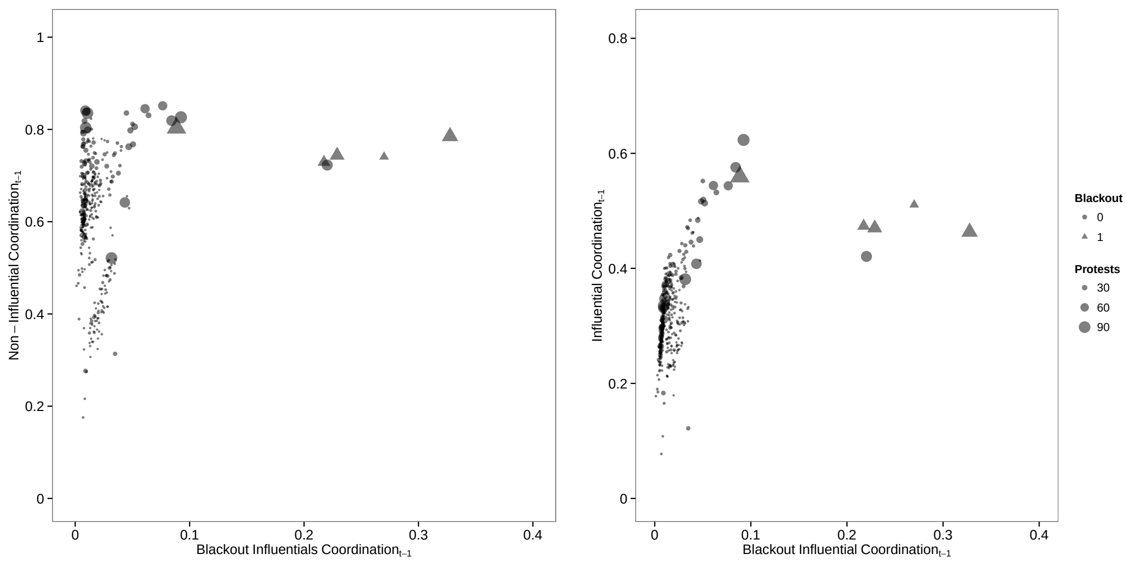
Egypt's January 25^{th} protests surprised everyone - activists, bystanders, and state authorities - with its large mobilization and brief occupation of Tahrir Square. The Mubarak regime had spent the previous days denying that the events in Tunisia would spread to their country, despite a spate of imitation immolations [Khalil, 2011, pg. 127]. Many Muslim Brotherhood leaders, despite not having sanctioned the protests, were summarily jailed, as the government assumed only it could mobilize such a crowd. With the next major protest called for January 28^{th} after Friday prayers, the government suspended cell phone service and internet access just after midnight on January 28^{th} (the morning of the 28^{th}). The government appears to have figured that people would not protest if they could not communicate with one another; the plan backfired, as Egyptians had no way to communicate except by going outdoors. Instead, the blackout led to more protests [Hassanpour, 2014].

Contrary to conventional wisdom, digital communication was not completely severed. One internet service provider, Noor, functioned through the end of January 31^{st}; it provided connectivity for critical government functions, the Cairo stock exchange, and several international hotels [Glanz & Markoff, 2011]. Individuals who knew to go to hotels or who had friends with access to Noor could therefore use Twitter; even on the one day without any internet access, February 1^{st}, one could use landlines, dial-up modems, and Google's 'Speak to Tweet' service to get online [Gunning & Baron, 2013, pg. 286]. The blackout therefore increased the cost of accessing the internet, limiting it to those with expertise or social connections with those who still had access.

Anyone observed tweeting from Egypt between January 28^{th} through February 1^{st} can therefore reasonably be classified as an belonging to the core of Egypt's network, regardless of their Lotan et. al coding or number of followers. This dataset observed 740 accounts that used Twitter from Egypt during the blackout, with a maximum of 338 tweeting on February 1^{st}. In terms of Twitter behavior, they are most similar to Egyptian bloggers and activists or those in the 99^{th} percentile of the follower distribution. These users, who I call the *blackout core*, have an average of 8046 followers, and are responsible for an average of 650 tweets and 12.05% of all tweets. How they use Twitter differs, however: they retweet less often (4.74% of their tweets are retweets) than bloggers and activists (8% and 7%) but about as often as those in the 99^{th} percentile, and they mention other users very infrequently - at 21.66%, less frequently than any other group in the sample. They use hashtags less frequently, in 21.66% of tweets, than bloggers, activists, or those in the 99^{th} percentile. Yet 59.36% of their tweets contain a link, more than any other Egyptian group except Mainstream Media. That the blackout identification accords with the follower-based measure of influence used throughout the paper provides reassurance about the validity of those measures.

Figure 1.6 shows how the blackout accounts' coordination correlates with $Coordination_{i,t}$ and $Core\ Coordination_{i,t}$ and protest (size of each point). $Blackout\ Core\ Coordination_{t}$ was calculated the same way as $Core\ Coordination_{i,t}$ except that having tweeted during the blackout, not number of followers, is the grouping variable. For both measures, shown respectively in Figure 1.6a and Figure 1.6b, there is strong correlation in the early part of new measure and the other two coordination variables. The day with the lowest level of lagged blackout coordination is January 28^{th}, the first day of the blackout; this result makes sense since January 27^{th} was a more representative sample of Egyptian Twitter users than later days would be. Just as with Figure 1.5.2, the key is to pay attention to the distribution of the size of the points: days with many protests occur across a range of $Blackout\ Coordination_{t}$ values.

Replicating the main models from Section 1.4 confirms that peripheral coordination drives protest mobilization. Table 1.8 shows these results: $Blackout\ Coordination_{i,t-1}$ is not significant in any model. The only stable result from the models is $Blackout\ Mention_{i,t-1}$, which is positive and significant. Though they infrequently mention other accounts, in comparison to the other categories used to delineate the core, they are more likely to do so during protest events.



(a) Blackout Influentials and Protest (b) Blackout, Non-Blackout Influentials

Figure 1.6: Blackout Influentials and Protest

1.9 Discussion

This paper claims that coordination occurs through those with few social connections, and this coordination leads to protest mobilization. These peripheral network individuals outweigh those in the core because protests diffuse through a complex contagion process, a process which, in the context of protests, requires distributed coordination to spread. These results join a growing body of quantitative work at the intersection of information and communication technology and state repression. Jan Pierskalla and Florian Hollenbach find that, in Africa, cell phone coverage increases the probability of violent conflict [Pierskalla & Hollenbach, 2013]. Jacob Shapiro and Nils Weidmann find the opposite effect in Iraq; using time-variant data on new cell phone coverage, they find that the provision of cellular coverage decreases insurgent violence [Shapiro & Weidmann, 2011]. Gary King, Jennifer Pan, and Molly Roberts measure censorship on Chinese blogs; they find that Chinese censors target posts

Table 1.8: Blackout Accounts do not Provide Coordination

	$Protest_t$				
	(1)	(2)	(3)	(4)	(5)
$\text{Coordination}_{t-1}$	0.592	4.420***	4.677***	4.283***	4.904***
	(0.447)	(0.952)	(1.037)	(1.036)	(1.176)
Hashtag $\%_{t-1}$		−11.399***			−10.446***
		(2.550)			(2.944)
Retweet $\%_{t-1}$		14.286*			13.877
		(7.767)			(10.516)
Link $\%_{t-1}$		−2.863			2.556
		(2.606)			(3.033)
Mention $\%_{t-1}$		−7.604**			−8.928***
		(3.086)			(3.363)
Protest_{t-1}	0.049***	0.034***	0.043***	0.046***	0.043***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.008)
Repression_{t-1}	0.012	0.014	0.011	0.017	0.008
	(0.018)	(0.018)	(0.018)	(0.018)	(0.017)
Blackout $\text{Coordination}_{t-1}$			−63.086***	−59.331***	−44.432**
			(17.901)	(22.485)	(22.651)
Blackout Hashtag $\%_{t-1}$			45.164***	22.291	19.221
			(13.042)	(18.716)	(19.351)
Blackout Retweet $\%_{t-1}$				50.241**	28.044
				(22.693)	(26.984)
Blackout Link $\%_{t-1}$				−3.888	−19.470***
				(4.163)	(5.792)
Blackout Mention $\%_{t-1}$				18.620**	31.304***
				(7.738)	(8.516)
Constant	0.776***	4.493*	−1.942***	−1.721**	3.390
	(0.263)	(2.299)	(0.692)	(0.698)	(2.488)
Country FE	No	No	No	No	No
N	425	415	415	415	415
Log Likelihood	−1,130.493	−1,080.934	−1,088.982	−1,080.382	−1,070.603

*p < .1; **p < .05; ***p < .01

which could generate collective action but are more permissive of writings critical of the Communist Party [King, Pan & Roberts, 2013].

Focusing on modern authoritarian countries may restrict the generalizability of this finding. It may also be that modern telecommunication technologies - phones, fax machines, and mass media, not just social media - facilitates mobilization by makes it easier for peripheral members to learn of other protests or coordinate without core social network actors [Weyland, 2012, Hale, 2013, Warren, 2014]; if that is the case, core social network members may have driven mobilization in earlier centuries. In countries where the state engages in less repression, especially before the start of protests, it may be that core social network members mobilize the periphery. Core network members may also behave in ways not recorded in events or social media data. Each possibility represents an intriguing avenue for subsequent research.

This paper uses Twitter data to make claims about off-Twitter behavior, but the usage of Twitter data introduces two sampling concerns. First, services exist which allow individuals to buy followers, so those accounts here identified as belonging to the core may therefore not be true core members. In fact, the accounts most likely to buy followers are those desiring to be in the core and so behaving like core members: politicians, celebrities, or small businesses [Stringhini et al., 2013]. Moreover, these types of accounts are likely to engage in more tweet and hashtag production than others, increasing the amount of coordination coming from the core [Wu et al., 2011]. While the buying of followers does not appear to have been a systematic practice in any of the countries during this study's period, that practice would bias in favor of finding core coordination.

Second, it is possible that individuals in the offline core select out of the Twitter core to avoid state repression, so the peripheral coordination may derive from peripheral Twitter accounts that belong to core members outside of Twitter. Though possible, work that has manually identified accounts in Tunisia and Egypt has been able to identify accounts belonging to political activists, suggesting online evasion may not be extensive [Lotan et al., 2011]. Identification of users from Lotan et. al (2011) in these data show that their tweets correlate negatively with subsequent protest (see Table 1.5, Model 1). Ethnographic work from Egypt shows that certain individuals believed their social network accounts to be monitored, but they did not respond by creating more social media accounts [Gunning & Baron, 2013, pgs. 284-287].

Despite the reliance on social media data, this paper does not address whether they, or telecommunications more broadly, affect protest. On one hand, social media may increase subsequent protest if it causes more individuals to learn about the state's actions and those individuals protest when they would not have without the knowledge-providing role of social media. Yet the knowledge-providing role could have counterbalancing effects: as more people learn the resolve of the state against protesters, fewer individuals may protest than otherwise would have. Appropriately answering this question requires data with very precise location information, preferably with temporal variation of social media presence. These data exist and have been used to test violence in Africa [Pierskalla & Hollenbach, 2013] and Iraq [Shapiro & Weidmann, 2011, Shapiro & Siegel, 2015], but the results are contradictory.

The paper assumes that behavior on online networks parallels that of offline interpersonal ones, allowing researchers to make inferences on heretofore hidden behaviors; research comparing behavior on Twitter to behavior off it lend credence to this assumption. Large analyzes of Twitter show that user behaviors there are the same as those observed offline: the distribution of followers follows a power-law distribution [Kwak et al., 2010], users connect to people who look like them (homophily) [Zamal, Liu & Ruths, 2012*a*], friends offline follow each other on Twitter [Xie et al., 2012], Dunbar's number applies to users on Twitter [Dunbar et al., 2015], and interaction decreases with geographic distance [Kulshrestha et al., 2012]. Substantively, survey work on activists in the U.S. Civil Rights movement and East Germany's 1989 protests shows that knowing others who had protested is the strongest correlate with a respondent's decision to participate [McAdam, 1986, Opp & Gern, 1993], and surveys of protesters in Tahrir Square show that individuals learned about the protests primarily through interpersonal connections or satellite television [Tufekci & Wilson, 2012]. Using Twitter data allows the researcher to observe these patterns across more spaces and time than before, but using the data does not reveal whether Twitter, and online social networks more generally, affects offline interpersonal networks.

No work has been able to show if social media *cause* protest, as it is very difficult to know which countries or regions of countries do or do not have a social media platform and then compare those areas to similar places without social media. Because of the difficulty of isolating social media's affect, this paper has chosen not to ask that question. The point of using social media data is to better understand our world. Social media data, especially that which is publicly available, resolves the temporal resolution problem facing previous work, but connecting those data with detailed spatial data is still a challenge. Because of limits in the data for protests and the paucity of tweets from these 16 countries with GPS coordinates, for example, analysis here was restricted to the country level.

This paper demonstrates the contributions big data can make to understanding processes of of social influence in social networks. Researchers have begun to understand how these data can provide new insights into political phenomena such as voting [Bond et al., 2012] or ideological sorting [Barbera, 2015]. These data primarily come from online social networks such as Facebook or Twitter, though anonymized call records, YouTube, or discussion boards [Nielsen, 2012] are often used. While an ideal research design would randomly assign treatments in order to measure influence

[Aral & Walker, 2012, Muchnik, Aral & Taylor, 2013], observational studies on these new datasets may allow researchers to observe and measure otherwise hidden mechanisms [Grimmer, 2015, Monroe et al., 2015, Shah, Cappella & Neuman, 2015]. For example, studies of protest mobilization have relied on *post hoc* surveys, but social media allow one to observe how individuals behaved, not how individuals reported they behaved.

Properly used, social media data should become another tool for researchers, but it is most likely to generate knowledge when used as a window into already existing processes [Bennett & Segerberg, 2013]. It is not clear that social media create new behaviors or fundamentally change social relations. Its main effect is to lower the cost of communication, and lowering the cost of communication also lowers the cost of data gathering. But lower costs do not clearly favor one group of actors over another: the printing press created Martin Luther's *95 Theses* and Russia's *Pravda*, and states have learned how to use the internet and social media to repress [Hoffman, 2015, Rod & Weidmann, 2015]. Using social media data to understand social behavior is therefore the main benefit of "big data". If social scientists have been stuck looking for keys under a streetlight, they now have access to stadium lights. Even stadium lights leave much of the world in the shadows.

1.10 Supplementary Materials

1.10.1 Twitter Data

The tweets involved in this analysis were extracted from Twitter's 10% API, an unbiased sample of 10% of all public activity on the platform. There are two ways in which country of origin was identified. First, if a Twitter user has enabled location sharing, the tweet will have GPS coordinates, and those coordinates are used to assign country-location for that tweet. If the country-location is one of the 16, the tweet is saved. Second, users can report their location as part of their profile, and that location is reported as metadata with each tweet. The user-reported location is then compared to a dictionary of cities and country names to assign each tweet to a city or country. For more detail, see the Materials and Methods section of Mocanu et. al 2013.

Table 1.9 shows 4 fictional tweets. The first shows what a user mention is; the second, a link; the third, a retweet; and the fourth, what a hashtag is and how it is used.

1.10.2 Construction of Coordination Measure

This section graphically illustrates the construction of the measure of coordination, the Gini coefficient for hashtag use on each country-day, as well as what hashtag usage looks like on low and high coordination days.

Figure 1.7 demonstrates the construction of the Gini coefficient. The outcome of interest, the number of times each hashtag is used on a day in a country, is ordered from least to most frequent (the hashtag(s) used once to the hashtag used the

Table 1.9: Tweet Typology

Mention	“@ZacharyST you’ll love the new spielberg movie”
Link	“What I watched last night http://bit.ly/hPHQG8”
Retweet	“RT @KingJames: We’ll get the Warriors next year”
Hashtag	“Will #decaprio finally get his #oscar?”

most). This arrangement generates a distribution similar to the ordered exponential probability distribution function (PDF) shown in Figure 1.7a. The observed PDF is converted to a cumulative distribution function (CDF), as shown in Figure 1.7b. The observed CDF is compared to what the CDF would look like if each hashtag was used the same number of times on a country-day, which is the ordered uniform CDF shown in Figure 1.7b. Coordination is then the ratio of the area between the ordered uniform CDF and the observed CDF to the area under the ordered uniform CDF, as shown in Figure 1.7c. A value of 1 means that all of the uses of a hashtag on a given day came from one hashtag, while a value of 0 means that all hashtags were used an equal number of times.

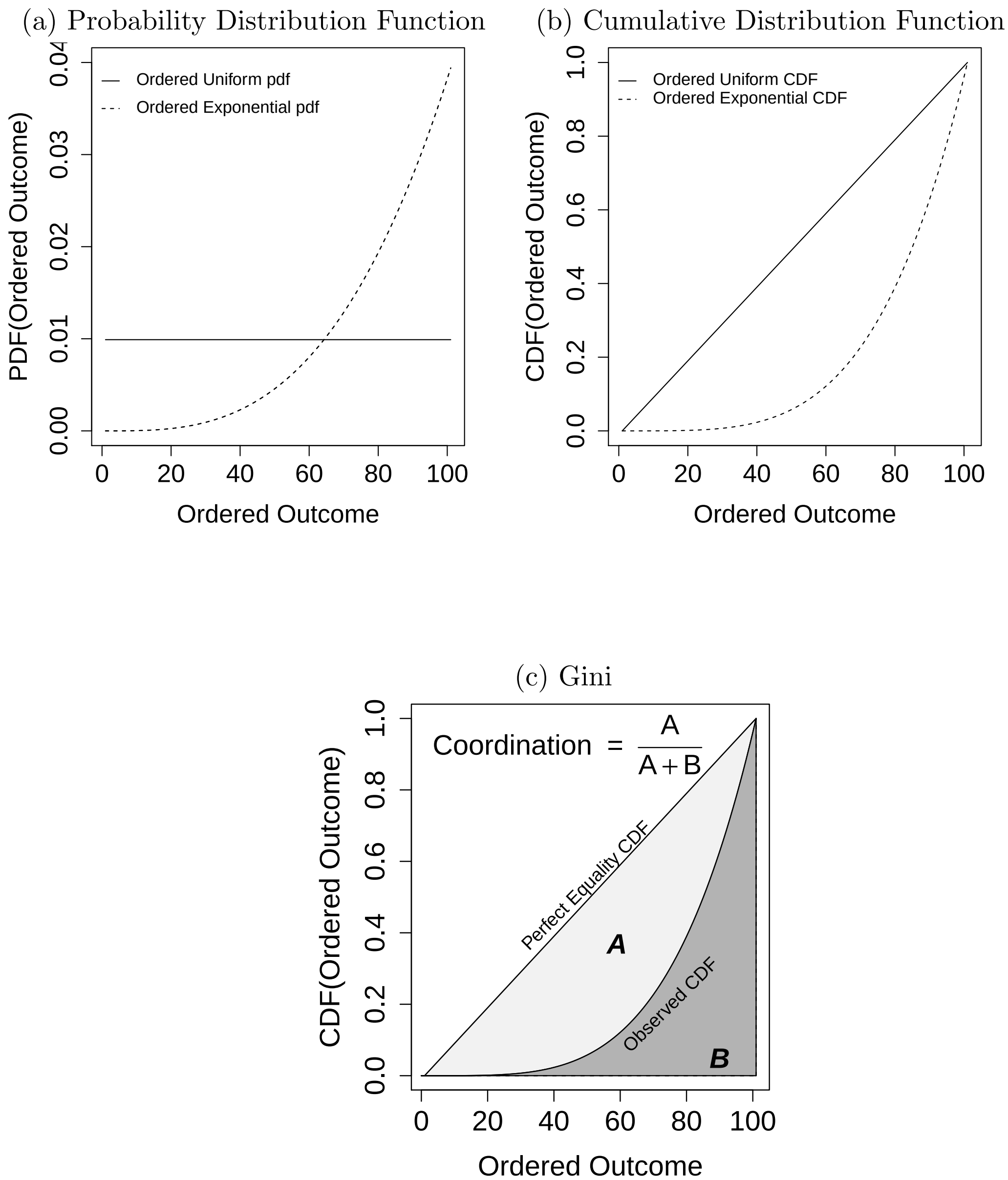


Figure 1.7: PDF $\rightarrow$ CDF $\rightarrow$ Gini

Figures 1.8, 1.9, and 1.10 show how the coordination measure was created for 3 days in the data. Figure 1.8 shows the day with the lowest coordination in this dataset, February 8^{th}, 2011 in Bahrain. Notice that the slope of the line is constant for the bottom 85% of hashtags, meaning most hashtags were used the same number of times (1). Figure 1.9 is the day with the most coordination in the dataset, February 2^{nd}, 2011 in Egypt. Notice here that the vast majority of the uses of a hashtag on that day, 1,867 out of 2,614 (71.42%), concentrate on the 3 most common hashtags, and 2 of the 3 are clearly political; #egypt was also used for protest communication at this time, though that is less apparent than #jan25 or #tahrir. Figure 1.10 shows the amount of coordination on January 24^{th}, 2011, just before Egypt's first major protest.

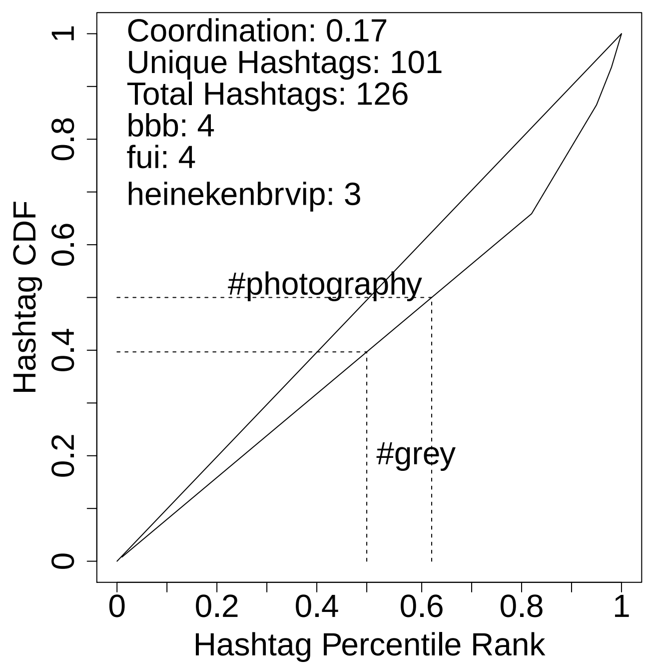


Figure 1.8: Least Coordination: Bahrain on 02.08.2011

Figure 1.11 examines *Core* $Coordination_{i,t}$ in more detail. The first panel shows that the lagged measure correlates highly with subsequent protest, much like $Coordination_{i,t-1}$; this result is not surprising, as the two variables have a .882 correlation. Figure 1.11b breaks *Core* $Coordination_{i,t}$ into its constituent parts, $Coordination_{i,t}$ and the percent of tweets with hashtags that are from influentials. In that figure, each day is sized by the number of protests.

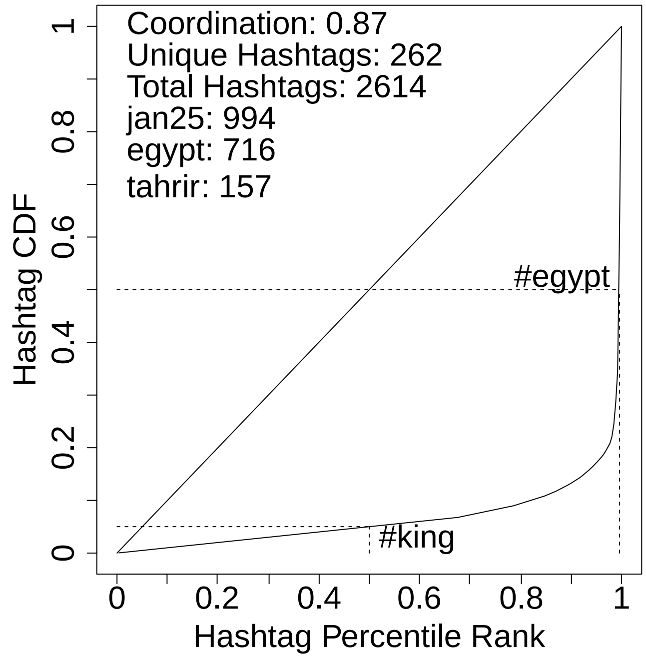


Figure 1.9: Most Coordination: Egypt on 02.02.2011

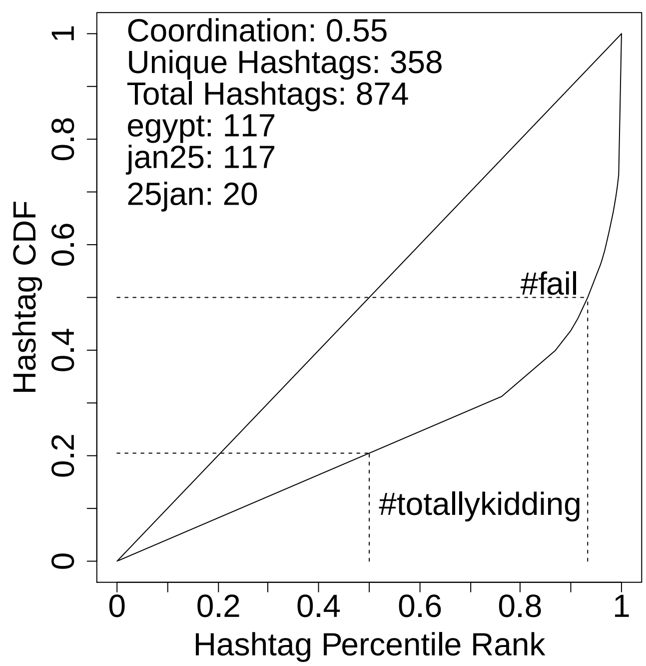
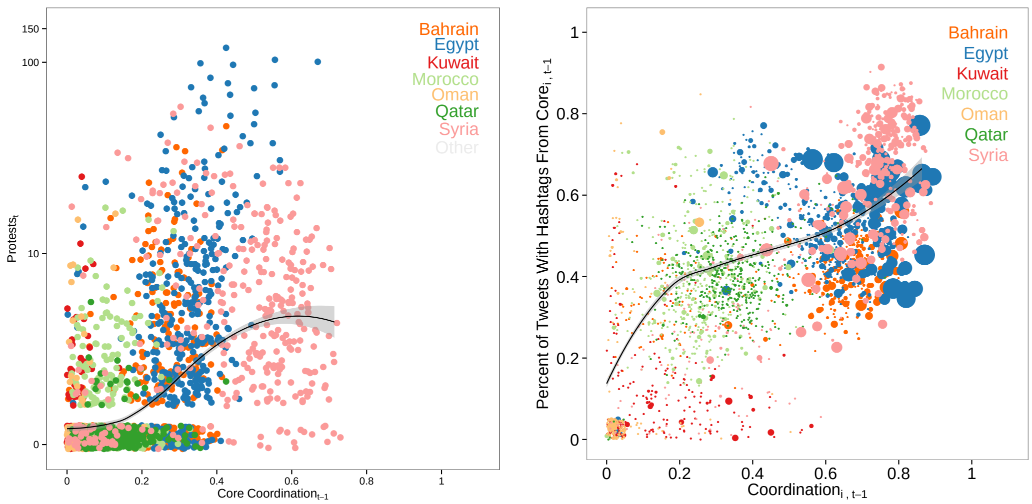


Figure 1.10: Egypt on 01.24.2011

How is it that the strong correlations shown in Figures **??** and **??** have disappeared? Figure 1.12 suggests why *Core Coordination*$_{i,t-1}$ does not correlate with

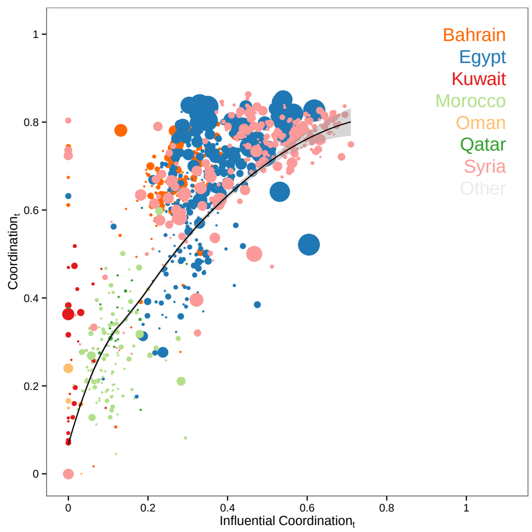


(a) Core Coord. and Protest (b) Constituent Parts of Core Coord.

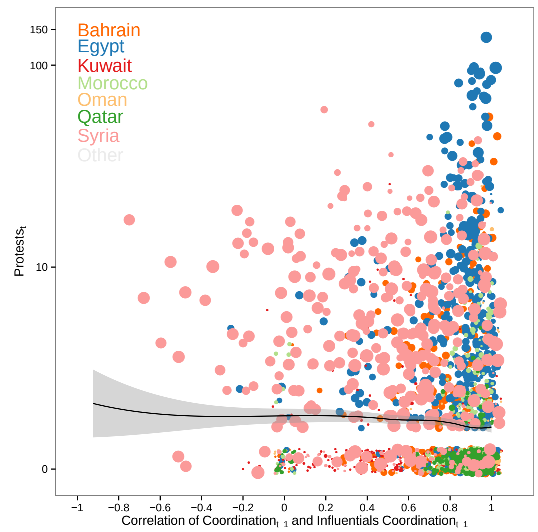
Figure 1.11: Core Coordination Variable

protest in a statistically meaningful way. The key insight is that the size of protest does not vary with *Core Coordination*$_{i,t-1}$. Figure 1.12a reveals the strong collinearity between the two coordination variables, but the distribution of the number of protests (size of points) is stable over most of the range of *Core Coordination*$_{i,t-1}$. In other words, the size of the dots (number of protests) increases from left to right, as coordination increases, but they increase much less as *Core Hashtag %*$_{i,t-1}$ increases. *Core Hashtag %*$_{i,t-1}$ is how *Core Coordination*$_{i,t-1}$ separates itself from *Coordination*$_{i,t-1}$, but it aids little in distinguishing levels of protest. Figure 1.12b shows how the correlation of the two coordination variables itself correlates with protest. As the two measures become more correlated, the number of protests decreases, though Syria drives this result. Looking just at the x-axis right of 0, there is no correlation between this measure and protest size. Figure 1.12c shows that there is similarly no correlation, once Syria is excluded, between the coordination variables' correlation and *Core Coordination*$_{i,t-1}$. There are days with many protests across the range of

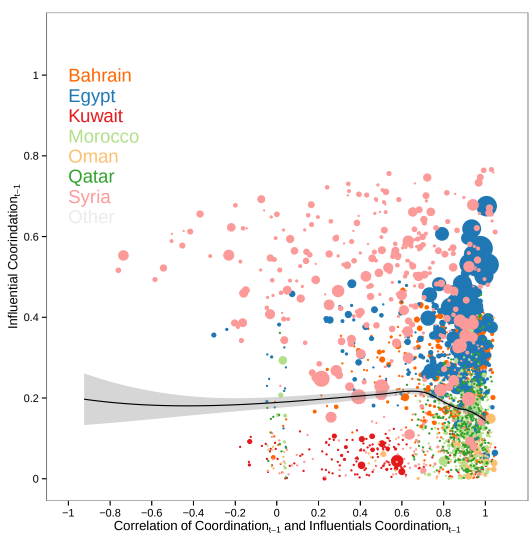
Core Coordination$_{i,t-1}$. Finally, notice the shape of the curve in Figure 1.11a: there is already little correlation between *Core Coordinationi*, $t-1$ and the next day's protest at high values of the variable. While *Coordination*$_{i,t-1}$ and *Core Coordination*$_{i,t-1}$ are highly correlated, that correlation appears to be driven entirely by the coordination occurring in the periphery.



(a) Sized by Protest



(b) Sized by Core Coordination



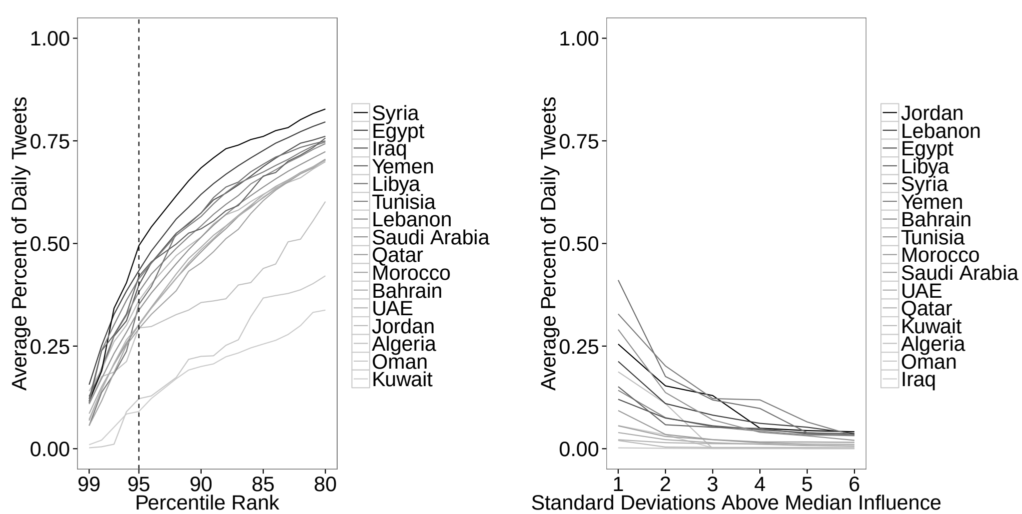
(c) Sized by Protest

Figure 1.12: Core Coordination Variable

1.10.3 Alternate Measures of Elites

The following figures show how tweet production and follower count vary depending on the threshold used.

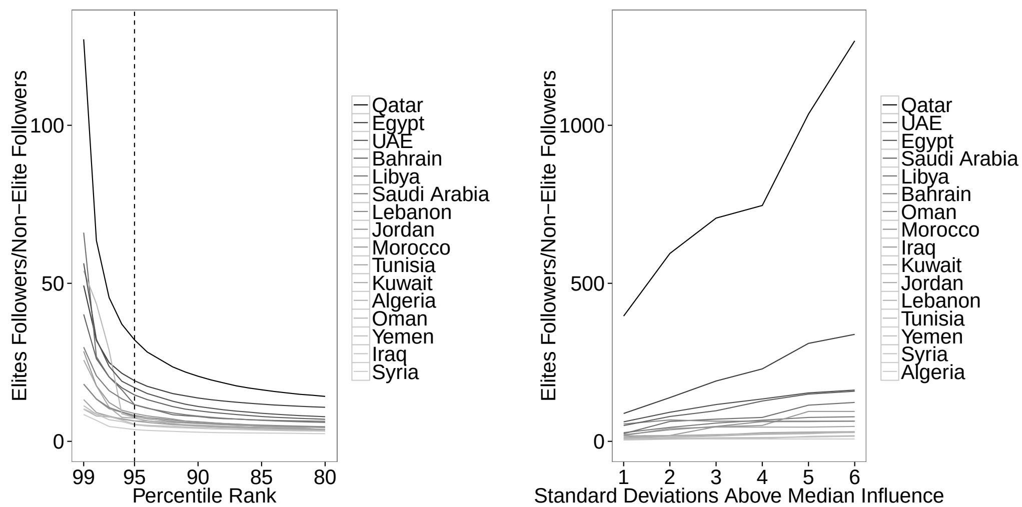
Figure 1.13 shows the average percent of each country's daily tweets which come from popular users, based on two measures. Figure 1.13a shows users in the 99th percentile, 98th percentile, and so on to the 80th percentile. The vertical dashed line represents the 95th percentile, the threshold used in the paper's regressions. Figure 1.14b uses standard deviations above each country's median number of followers as its threshold. The lines now decrease because each 1 unit increase in standard deviation creates a smaller group, so there have to be fewer tweets in that group than the one immediately proceeding. The legends are ordered based on the 80th percentile: the country with the highest average percentage of daily tweets coming from users at or above the 80th percentile is Syria, then Egypt, so on down to Kuwait.



(a) Percentile (b) Standard Deviation Rank

Figure 1.13: Daily Average Tweet Production by Popularity Threshold

Figure 1.14 is similar to Figure 1.13 except the outcome measure is the ratio of followers for core users to followers for peripheral ones. Figure 1.14a sorts users based on percentiles, while Figure 1.14b sorts on standard deviations above median number of followers. In Qatar, users in the 99th percentile based on number of followers have approximately 150 times as many followers as a Qatari with the median number of followers; users with 1 standard deviation more followers than the median user have approximately 400 times as many followers as the median Qatari. As in Figure 1.13, the legend for each subfigure is ordered from highest to lowest.



(a) Percentile (b) Standard Deviation Rank

Figure 1.14: Elite Follower Ratio by Popularity Threshold

The takeaway from these charts is that there is no inexpensive way to identify core members that will find the same kind of accounts across countries. The optimal approach to validating the chosen measure is therefore to run the main model separately as many times as there are core threshold measures. The results from this robustness check are reported in the paper's body text.

1.10.4 Exploring Upper Tail of Core Distribution

Table 1.10 shows descriptive statistics for core users. It replicates the presentation in Table 1 of the main paper. A few interesting trends emerge from Table 1.10. First, notice how few accounts are in the very upper tail (99^{th} percentile or above, or 6 standard deviations above the median number of followers) of each country. Many countries have fewer than 10 users in these bins, and those users often appear in the dataset only a few times; see Oman, Yemen, and Algeria are in this bin. Second, there exists a steep descent in the number of followers from the 99.9^{th} percentile to the 99^{th}, and the descent is almost as steep down to the 95^{th} percentile. Third, notice that the 99^{th} and 99.9^{th} percentiles use hashtags and links at much higher rates than those even slightly below them in their country's followers' distribution. Fourth, and finally, notice that some accounts identified as being Tunisian or Egyptian appear in other countries as well, such as Lebanon, Morocco, and Qatar. This finding was unexpected and worth future investigation.

Table 1.10: Core Threshold Statistics by Country

Algeria	1 SD Above Med.	15	1621.01	93.33	0.27	0.43	0.59	0.79
Algeria	6 SD Above Med.	2	6786.50	1.50	0.33	0.00	0.33	1.00
Algeria	99.9th Percentile	1	9511.00	1.00	1.00	0.00	1.00	1.00
Algeria	99th Percentile	5	3582.28	164.20	0.10	0.72	0.73	0.97
Algeria	Core 95 Percentile	23	1183.08	95.61	0.35	0.28	0.39	0.60
Bahrain	1 SD Above Med.	365	4445.40	412.00	0.36	0.04	0.44	0.36
Bahrain	6 SD Above Med.	64	14938.11	523.58	0.17	0.03	0.56	0.50
Bahrain	99.9th Percentile	21	25657.88	824.91	0.09	0.02	0.58	0.58
Bahrain	99th Percentile	201	6873.44	449.57	0.29	0.03	0.50	0.42
Bahrain	Core 95 Percentile	1005	1995.99	319.24	0.42	0.05	0.38	0.28
Egypt	Mainstream Media	1	103927.00	5281.00	0.00	0.00	0.74	0.70
Egypt	Non-Media Org.	2	23877.40	457.50	0.32	0.07	0.22	0.53
Egypt	MSM Employee	9	22463.50	650.22	0.41	0.01	0.21	0.21
Egypt	Blogger	15	8394.17	1070.67	0.52	0.08	0.33	0.22
Egypt	Activist	10	8036.60	703.40	0.42	0.07	0.28	0.33
Egypt	1 SD Above Med.	543	9585.76	830.09	0.43	0.06	0.27	0.34
Egypt	6 SD Above Med.	104	31542.21	1295.48	0.26	0.05	0.37	0.57
Egypt	99.9th Percentile	80	37001.28	924.69	0.33	0.02	0.39	0.44
Egypt	99th Percentile	793	7104.31	736.38	0.46	0.05	0.27	0.32
Egypt	Core 95 Percentile	3962	1868.80	409.94	0.45	0.04	0.25	0.30
Iraq	1 SD Above Med.	17	19710.78	485.24	0.62	0.06	0.39	0.09
Iraq	6 SD Above Med.	1	236675.00	1.00	0.00	0.00	0.00	1.00
Iraq	99.9th Percentile	5	57288.40	59.60	0.16	0.06	0.75	0.11
Iraq	99th Percentile	45	8576.75	414.38	0.45	0.20	0.25	0.10

Table 1.10 Continued: Core Threshold Statistics by Country

Country	Group	Accounts	Followers Avg.	Tweet Avg.	Mention %	Retweet %	Hashtag %	Link %
Iraq	Core 95 Percentile	224	2254.30	260.17	0.40	0.14	0.20	0.20
Jordan	1 SD Above Med.	251	2117.48	277.47	0.37	0.06	0.24	0.51
Jordan	6 SD Above Med.	24	9621.04	470.92	0.19	0.03	0.35	0.73
Jordan	99.9^{th} Percentile	9	15371.72	282.11	0.44	0.03	0.12	0.44
Jordan	99^{th} Percentile	86	4342.62	446.14	0.24	0.03	0.20	0.71
Jordan	Core 95 Percentile	426	1465.35	231.61	0.39	0.05	0.24	0.49
Kuwait	1 SD Above Med.	56	7332.00	10.64	0.59	0.01	0.08	0.14
Kuwait	6 SD Above Med.	10	25240.53	7.00	0.40	0.07	0.07	0.30
Kuwait	99.9^{th} Percentile	4	43236.25	6.75	0.78	0.00	0.07	0.04
Kuwait	99^{th} Percentile	32	11216.00	9.09	0.47	0.02	0.10	0.13
Kuwait	Core 95 Percentile	158	3344.44	17.13	0.65	0.01	0.09	0.09
Lebanon	MSM Employee	1	56753.60	5.00	0.60	0.00	0.00	0.40
Lebanon	Blogger	2	7311.00	1.00	0.00	0.00	0.50	0.00
Lebanon	Activist	1	6681.62	170.00	0.74	0.01	0.11	0.15
Lebanon	1 SD Above Med.	368	2381.27	302.15	0.37	0.10	0.27	0.41
Lebanon	6 SD Above Med.	42	9173.22	458.29	0.18	0.02	0.25	0.67
Lebanon	99.9^{th} Percentile	18	14186.87	385.39	0.19	0.01	0.36	0.51
Lebanon	99^{th} Percentile	173	3886.31	340.76	0.28	0.12	0.29	0.52
Lebanon	Core 95 Percentile	863	1306.22	231.43	0.41	0.09	0.28	0.37
Libya	Blogger	1	2147.00	2.00	0.00	0.00	1.00	0.00
Libya	1 SD Above Med.	79	4910.78	162.68	0.47	0.04	0.36	0.19
Libya	6 SD Above Med.	10	18325.41	293.90	0.39	0.05	0.60	0.26
Libya	99.9^{th} Percentile	4	31026.75	369.00	0.20	0.07	0.80	0.34

Table 1.10 Continued: Core Threshold Statistics by Country

Country	Group	Accounts	Followers Avg.	Tweet Avg.	Mention %	Retweet %	Hashtag %	Link %
Libya	99^{th} Percentile	37	8482.10	129.11	0.43	0.06	0.47	0.23
Libya	Core 95 Percentile	182	2540.17	164.10	0.35	0.19	0.31	0.38
Morocco	Blogger	1	1671.32	244.00	0.53	0.38	0.23	0.36
Morocco	1 SD Above Med.	142	6061.31	196.80	0.50	0.13	0.22	0.37
Morocco	6 SD Above Med.	22	23346.91	212.50	0.45	0.03	0.28	0.36
Morocco	99.9^{th} Percentile	10	37980.35	159.40	0.29	0.01	0.39	0.57
Morocco	99^{th} Percentile	98	8022.34	172.94	0.43	0.11	0.22	0.45
Morocco	Core 95 Percentile	489	2380.80	179.68	0.48	0.09	0.24	0.35
Oman	1 SD Above Med.	6	47661.06	3.00	0.33	0.00	0.44	0.17
Oman	6 SD Above Med.	2	122889.50	1.50	0.677	0.00	0.00	0.33
Oman	99.9^{th} Percentile	1	130487.00	2.00	1.00	0.00	0.00	0.00
Oman	99^{th} Percentile	7	41600.91	2.71	0.32	0.00	0.42	0.21
Oman	Core 95 Percentile	33	10328.22	31.42	0.55	0.03	0.11	0.32
Qatar	Mainstream Media	2	281861.00	1381.00	0.08	0.00	0.50	0.91
Qatar	MSM Employee	3	6360.29	836.67	0.65	0.01	0.27	0.19
Qatar	Blogger	1	2838.00	1.00	1.00	0.00	1.00	0.00
Qatar	1 SD Above Med.	58	26021.46	417.91	0.41	0.06	0.30	0.34
Qatar	6 SD Above Med.	8	127181.55	812.63	0.15	0.00	0.40	0.62
Qatar	99.9^{th} Percentile	22	56340.09	634.95	0.36	0.05	0.29	0.39
Qatar	99^{th} Percentile	217	8590.11	349.83	0.43	0.05	0.21	0.21
Qatar	Core 95 Percentile	1083	2218.35	307.21	0.46	0.06	0.16	0.12
Saudi Arabia	MSM Employee	2	53017.31	767.00	0.79	0.00	0.08	0.08
Saudi Arabia	1 SD Above Med.	425	9345.20	580.17	0.56	0.03	0.24	0.28

Table 1.10 Continued: Core Threshold Statistics by Country

Country	Group	Accounts	Followers Avg.	Tweet Avg.	Mention %	Retweet %	Hashtag %	Link %
Saudi Arabia	6 SD Above Med.	74	32760.58	631.07	0.67	0.02	0.10	0.18
Saudi Arabia	99.9^{th} Percentile	89	28922.06	576.03	0.67	0.02	0.11	0.18
Saudi Arabia	99^{th} Percentile	885	5243.44	544.22	0.57	0.03	0.21	0.24
Saudi Arabia	Core 95 Percentile	4425	1448.05	337.76	0.58	0.04	0.14	0.17
Syria	1 SD Above Med.	148	1295.07	637.03	0.16	0.05	0.59	0.73
Syria	6 SD Above Med.	13	4891.79	598.00	0.15	0.02	0.44	0.72
Syria	99.9^{th} Percentile	4	1003.41	690.00	0.13	0.01	0.56	0.74
Syria	99^{th} Percentile	37	2927.40	719.65	0.21	0.02	0.56	0.67
Syria	Core 95 Percentile	182	1134.64	623.71	0.15	0.05	0.62	0.75
Tunisia	Mainstream Media	2	5604.50	741.00	0.16	0.12	0.77	0.79
Tunisia	MSM Employee	1	52503.00	1.00	0.00	0.00	0.00	1.00
Tunisia	Blogger	3	1910.77	258.33	0.57	0.13	0.30	0.20
Tunisia	Activist	4	2496.28	57.00	0.60	0.09	0.36	0.29
Tunisia	1 SD Above Med.	172	2499.44	385.69	0.45	0.13	0.30	0.41
Tunisia	6 SD Above Med.	15	11473.11	311.87	0.51	0.06	0.16	0.51
Tunisia	99.9^{th} Percentile	7	17749.31	206.71	0.25	0.06	0.22	0.68
Tunisia	99^{th} Percentile	62	4880.14	410.92	0.37	0.06	0.28	0.55
Tunisia	Core 95 Percentile	307	1681.87	308.22	0.47	0.11	0.31	0.38
UAE	1 SD Above Med.	240	15377.37	251.14	0.33	0.05	0.31	0.41
UAE	6 SD Above Med.	30	71908.66	342.30	0.27	0.06	0.42	0.32
UAE	99.9^{th} Percentile	48	50588.01	343.21	0.32	0.06	0.37	0.31
UAE	99^{th} Percentile	471	9018.08	232.85	0.39	0.07	0.26	0.37
Yemen	1 SD Above Med.	55	1217.52	367.22	0.13	0.03	0.51	0.67

Table 1.10 Continued: Core Threshold Statistics by Country

Country	Group	Accounts	Followers Avg.	Tweet Avg.	Mention %	Retweet %	Hashtag %	Link %
Yemen	6 SD Above Med.	7	4405.43	280.00	0.06	0.01	0.33	0.72
Yemen	99.9^{th} Percentile	2	6694.44	452.00	0.01	0.00	0.01	0.83
Yemen	99^{th} Percentile	14	2868.15	532.21	0.10	0.02	0.60	0.69
Yemen	Core 95 Percentile	69	1039.71	374.57	0.13	0.03	0.52	0.70

1.10.5 Potential Model Misspecification

The main model may not capture the correct data generating process, so different specifications, in addition to those in the verification section of the main paper, are presented below. Table 1.11 shows these and verifies that likely confounds have not driven the results. In the first column, every variable is lagged by 3 time periods, though only results for the one-day lags are shown. $Coordination_{i,t-1}$ is still positive and significant but not 2 or 3 days before a protest. None of the variables of core members' activity, even the lags, are significant, further showing that their behavior does not affect subsequent protest.

Table 1.11: Robust to Model Specification

	$Protest_{i,t}$		
	3 Lags	Only Arabic	Friday Fixed Effects
	(1)	(2)	(3)
$\text{Coordination}_{i,t-1}$	1.454**	2.324***	2.518***
	(0.593)	(0.502)	(0.657)
Hashtag $\%_{i,t-1}$	0.399	0.877**	0.541
	(0.692)	(0.379)	(0.638)
Retweet $\%_{i,t-1}$	−0.062	1.422***	−0.214
	(1.077)	(0.506)	(1.027)
Link $\%_{i,t-1}$	−0.656	−0.083	−0.768**
	(0.444)	(0.424)	(0.382)
Mention $\%_{i,t-1}$	−0.551	−0.315	−0.872**
	(0.481)	(0.515)	(0.390)
$\text{Repression}_{i,t-1}$	0.004	0.023*	0.019*
	(0.010)	(0.012)	(0.011)
$\text{Protest}_{i,t-1}$	0.099***	0.115***	0.119***
	(0.008)	(0.010)	(0.010)
Core Hashtag $\%_{i,t-1}$	0.626	0.580	0.930***
	(0.504)	(0.377)	(0.340)
Core Retweet $\%_{i,t-1}$	0.233	−0.153	0.142
	(0.375)	(0.251)	(0.385)
Core Link $\%_{i,t-1}$	0.187	0.176	0.685
	(0.521)	(0.565)	(0.577)
Core Mention $\%_{i,t-1}$	0.206	0.037	−0.240
	(0.347)	(0.201)	(0.268)
$\text{Friday}_{i,t}$			0.539***
			(0.182)
$\text{Coordination}_{i,t-1}$*Core Hashtag $\%_{i,t-1}$	−1.004	−1.136	−1.740*
	(1.391)	(0.842)	(0.980)
Intercept	−2.225***	−1.210***	−1.080***
	(0.589)	(0.330)	(0.207)
Country FE	Yes	Yes	Yes
N	6,468	6,397	6,620
Log Likelihood	−8,083.642	−8,074.174	−8,244.873

*p < .1; **p < .05; ***p < .01

It is also possible that English tweets drive the results. Twitter did not introduce an Arabic interface until March 2012 (though it did accept Arabic input for tweets), so users in the study may have needed a passing familiarity with English. Moreover, cellphones, especially smart phones, were not widespread in 2011, and an even smaller percentage of the population was on Twitter [Mourtada & Salem, 2011, Tufekci & Wilson, 2012, *International Telecommunications Union Statistics*, 2014]. Appealing to international actors is a common tactic of activists [Keck & Sikkink, 1998] and was a prominent use of Twitter by Arab Spring activists [Howard & Hussain, 2011, Howard et al., 2011, Aday et al., 2012, Starbird & Palen, 2012]. Since individuals in the protest countries are most likely not going to consume protest information in English, it is possible that international attention-seeking drives the results. To rule out this possibility, the main model is rerun but ignoring English tweets.[11] Replicating the main model, Column 2 in Table 1.11, using only Arabic tweets confirms the importance of peripherals' coordination and impotence of those in the core.

The main model does not account for potential periodicity in the data: Friday, the main religious day for Muslims, also saw surges in protest. In Egypt, for example, January 25^{th} was chosen as the first protest day because it coincided with National Police Day, a national holiday commemorating a massacre of Egyptian police by the British in 1952. January 25^{th} was a Tuesday, and many people returned home that night and to work the next day. Momentum then built for another major protest on January 28^{th}, a Friday, when individuals surged to Tahrir Square after the end of prayers [Ghonim, 2012].[12] To make sure that $Coordination_{i,t-1}$ does not capture religious effects, a fixed effect for Fridays is added. The model with Friday fixed effects, shown in Column 3 of Table 1.11, shows that coordination occurs independent of Friday effects, though Fridays do have more protests than non-Fridays. Column 3 also suggests some effect for the percent of hashtags originating in the core, though

[11] A Python implementation of Google's Compact Language Detector was used to identify each tweet's language.

[12] It is not a coincidence that the state shut-down all internet and phone service (save the internet connection for the Cairo stock exchange) late at night on the 27^{th}.

coordination from the core is not statistically significant and negative.

Table 1.12 shows that the results hold when countries at the extreme of the protest and tweet distribution are dropped. To identify countries to drop, each country's average number of protests and tweets per capita over the entire sample were calculated. Model 1 drops the 5 countries with the fewest protests per capita, Model 2 drops the 5 with the most. Models 3 and 4 do the same but with tweets per capita. In all models, peripheral coordination is positive and statistically significant while core coordination is negative and significant. The only behavior from the core that may be significant and positive is Core Hashtag $\%_{i,t-1}$, though the effect is out weighted by $Coordination_{i,t-1}$.

Table 1.12: Robust to Removing Countries

	Protest$_{i,t}$			
	Top 10 Protests	Bottom 10 Protests	Top 10 Tweets	Bottom 10 Tweets
	(1)	(2)	(3)	(4)
Coordination$_{i,t-1}$	2.908***	1.861**	3.373***	2.470***
	(0.627)	(0.815)	(0.802)	(0.554)
Hashtag %$_{i,t-1}$	0.682	0.090	0.898	0.452
	(0.658)	(0.715)	(0.598)	(0.615)
Retweet %$_{i,t-1}$	−0.071	−0.328	−1.785	−0.385
	(1.326)	(1.359)	(1.464)	(1.051)
Link %$_{i,t-1}$	−1.048**	−0.257	−1.403**	−0.860**
	(0.450)	(0.329)	(0.560)	(0.430)
Mention %$_{i,t-1}$	−0.933	−1.098***	−1.335**	−1.011**
	(0.570)	(0.381)	(0.640)	(0.400)
Repression$_{i,t-1}$	0.018	−0.005	0.013	0.027**
	(0.013)	(0.011)	(0.009)	(0.011)
Protest$_{i,t-1}$	0.106***	0.116***	0.090***	0.106***
	(0.010)	(0.015)	(0.010)	(0.010)
Core Hashtag %$_{i,t-1}$	0.708*	0.974*	1.595	0.907**
	(0.427)	(0.497)	(1.052)	(0.375)
Core Retweet %$_{i,t-1}$	0.060	0.758	−0.390	0.122
	(0.479)	(0.473)	(0.587)	(0.440)
Core Link %$_{i,t-1}$	1.138*	0.658	0.716	1.059**
	(0.602)	(0.723)	(0.815)	(0.517)
Core Mention %$_{i,t-1}$	0.007	−0.475*	0.190	−0.317
	(0.348)	(0.259)	(0.560)	(0.228)
Coordination$_{i,t-1}$*Core Hashtag %$_{i,t-1}$	−2.170**	−2.809*	−1.744*	−1.969*
	(1.016)	(1.644)	(0.969)	(1.005)
Intercept	−1.600***	−1.229***	−1.748***	−0.978***
	(0.331)	(0.263)	(0.503)	(0.240)
Country FE	Yes	Yes	Yes	Yes
N	4,551	4,545	4,559	4,545
Log Likelihood	−6,900.976	−4,944.046	−5,694.607	−6,674.396

*p < .1; **p < .05; ***p < .01

1.10.6 Placebo Tests

Two placebo tests confirm that the results on coordination are not a result of measurement error.

First, the coordination measure is calculated for each country-day while excluding the top 5, 10, and 20 most common hashtags for each day. The resulting coordination is therefore the coordination that occurs on less common hashtags, which are more likely to be non-protest hashtags. (Hashtags were not manually identified so that the measure could scale easily across countries and days.) If coordination on less common hashtags correlates with protest as much as coordination on all hashtags, then the operationalization has mistaken chatter focused on non-protest hashtags with protest coordination.

The results in Table 1.13 confirm that the original $Coordination_{i,t}$ measures protest coordination. The table shows the results of models where an increasing number of the most common hashtags per day - 5, 10, then 20 - are excluded from $\text{Coordination}_{i,t}$. The coefficient on $\text{Coordination}_{i,t}$ ranges from 39.3% to 43.6% smaller (it is 2.575 when not excluding hashtags) and decreases as an increasing number of hashtags are removed. At the same time, the coefficient on $Hashtag\ \%_{i,t-1}$ is now significant in each model (it was not significant in the original model), increases as more hashtags are removed, and almost triples is size compared to the full model. Not measuring the coordinating effect of the most popular hashtags pushes the correlation to $Hashtag\ \%_{i,t-1}$. Note as well that the signs, coefficient size, and results on the other coefficients are very similar to their values in the original model.

Table 1.13: Coordination in the Hashtag Long Tail

	$Protest_{i,t}$		
	Remove Top 5	Remove Top 10	Remove Top 20
	(1)	(2)	(3)
$\text{Coordination}_{i,t-1}$	1.563^{**}	1.537^{**}	1.450^{**}
	(0.687)	(0.686)	(0.721)
$\text{Hashtag } \%_{i,t-1}$	1.305^{*}	1.423^{*}	1.534^{**}
	(0.687)	(0.727)	(0.677)
$\text{Retweet } \%_{i,t-1}$	-0.388	-0.473	-0.505
	(0.967)	(0.980)	(0.952)
$\text{Link } \%_{i,t-1}$	-0.772^{**}	-0.771^{**}	-0.776^{**}
	(0.386)	(0.378)	(0.372)
$\text{Mention } \%_{i,t-1}$	-0.884^{**}	-0.889^{**}	-0.877^{**}
	(0.382)	(0.389)	(0.387)
$\text{Protest}_{i,t-1}$	0.121^{***}	0.122^{***}	0.122^{***}
	(0.011)	(0.011)	(0.011)
$\text{Repression}_{i,t-1}$	0.023^{*}	0.023^{*}	0.022^{*}
	(0.013)	(0.013)	(0.013)
$\text{Core Hashtag } \%_{i,t-1}$	0.666	0.692	0.687
	(0.512)	(0.512)	(0.510)
$\text{Core Retweet } \%_{i,t-1}$	0.244	0.249	0.234
	(0.451)	(0.461)	(0.456)
$\text{Core Link } \%_{i,t-1}$	0.632	0.633	0.639
	(0.549)	(0.539)	(0.535)
$\text{Core Mention } \%_{i,t-1}$	-0.259	-0.286	-0.303
	(0.247)	(0.242)	(0.238)
$\text{Coordination}_{i,t-1}\text{*Core Hashtag } \%_{i,t-1}$	-1.085	-1.185	-1.152
	(0.986)	(0.946)	(0.907)
Intercept	-0.953^{***}	-0.976^{***}	-0.988^{***}
	(0.214)	(0.197)	(0.183)
Country FE	Yes	Yes	Yes
N	6,620	6,620	6,620
Log Likelihood	$-8{,}303.888$	$-8{,}305.026$	$-8{,}306.342$

$^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$

Second, the main model, Model 5 from Table **??** in the main paper, is rerun with different lags and leads on $Coordination_{i,t-1}$; if $Coordination_{i,t-1}$ does not measure coordination, there should be no change in coefficient size as the lags and leads change. In fact, as Figure 1.15 shows, the effect size is much larger for a 1-day lag than it is for 2 or 3 day lags, and thereafter it decreases monotonically. The models suggest a positive correlation between a day's protest and future correlation, but this correlation is never as strong as it is for a 1-day lag, and it decreases as the future moves further away. These results are consistent with the theory of spontaneous collective action, as the theory does not say protest cannot affect future coordination, only that past coordination affects future protest. As Table 1.11 of the SM shows, controlling for past protests does not change the results on $Coordination_{i,t-1}$.

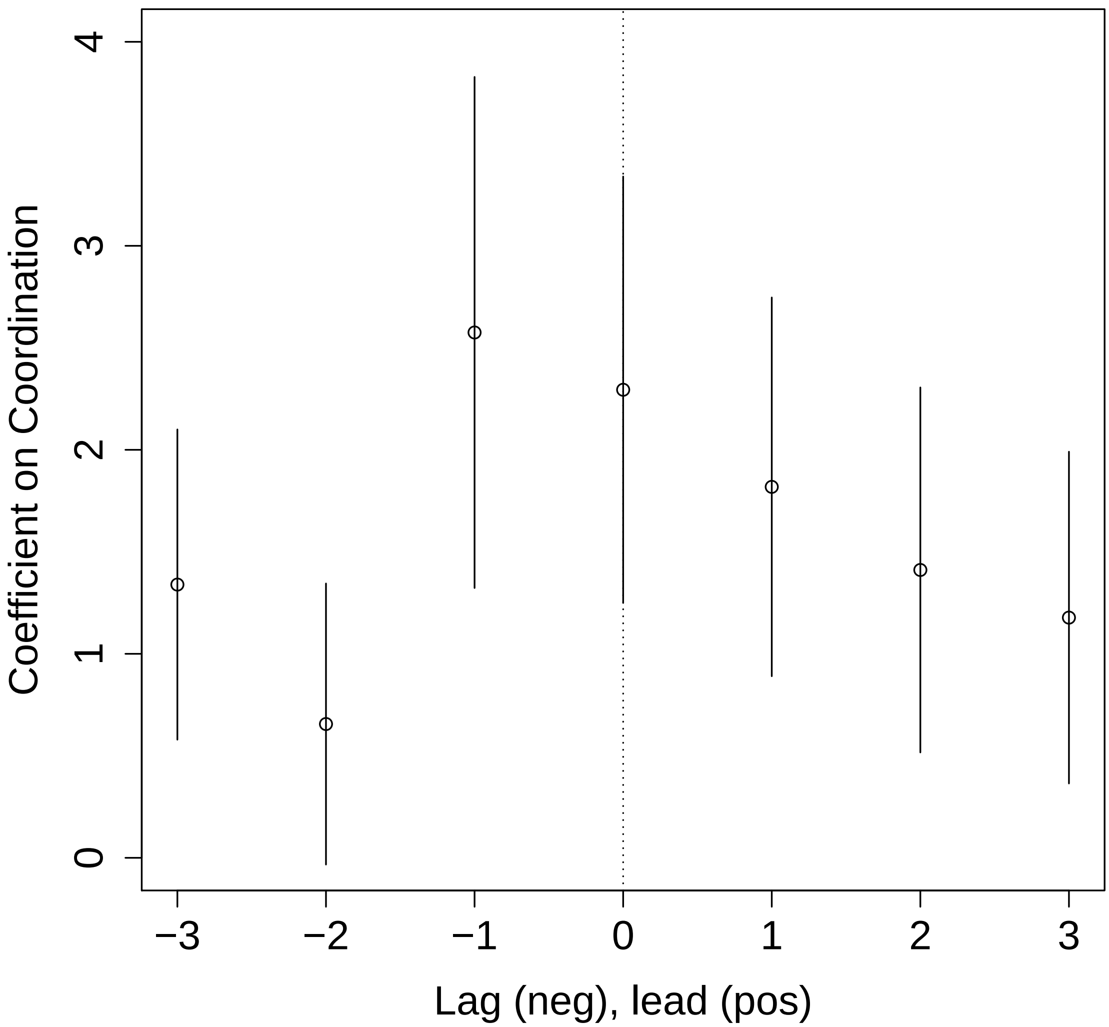


Figure 1.15: $Coordination_{i,t}$ Peaks with 1-Day Lag

1.10.7 Events Data

The overriding concern with using ICEWS, and machine-coded events data more generally, is measurement validity. This concern has two subcomponents. First, such data may capture news coverage of events more than it captures actual events, e.g. a large number of "protests" may actually mean a large number of articles about a single protest. Machine-coded events data, including ICEWS, may therefore not record the true number of protests. Second, news sources overreport novel events and underreport persistent ones (such as civil war or ongoing protests), biasing event counts even if duplication is not a problem [Davenport & Ball, 2002]. ICEWS' noise, though very low, may therefore be biased.

To the first concern: ICEWS correlates strongly with ACLED, a handcoded dataset, and the Global Database of Events, Location, and Tone (GDELT), another machine coded dataset. This correlation matches previous analyses. While no correlation statistics are given, time series figures and tables provided in an analysis from Michael Ward's team, the first academic users of ICEWS, shows that GDELT captures the same changes in activity as ICEWS [Ward et al., 2013]. Another analysis comparing the two shows that ICEWS correlates highly with GDELT and does best on conflictual events [Arva et al., 2013]. See the main paper for a discussion of that finding.

Hand-coded datasets have the highest levels of deduplication, as the dedepulication rules can be as flexible as necessary because of human oversight. A comparison of GDELT to two hand-coded datasets of violence in Africa finds that GDELT correlates with ACELD at .64, the Geo-Referenced Event Dataset at .33 [Hammond & Weidmann, 2014]; similar comparisons do not exist for ICEWS. The Social Conflict in Africa Dataset is another hand-coded dataset that tracks protest, among other events, across Africa (and Mexico, Central America, and the Caribbean) from 1990-2013 using handcoded articles from the *Associated Press* and *Agence France Presse* [Hendrix et al., 2012]. To compare SCAD to ICEWS, all events that were types 1 (organized demonstration), 2 (spontaneous demonstration), 3 (organized riot), or 4 (spontaneous riot) from Algeria, Egypt, Libya, Morocco, and Tunisia were selected. Those events are matched to their corresponding days in the ICEWS data, and only those events labeled as nationwide are kept. Because SCAD aggregates events which occur simultaneously in different parts of a country that are of the same type and topic, "nationwide" encapsulates more than one (but how many is unknown) event. The estimated number of participants, which is reported on a log scale, is then taken to see if it correlates with the number of protests recorded by ICEWS for that country-day. The results are shown in Figure 1.16. The more attendees at nationwide protests, the more protests ICEWS records.[13]

GDELT correlates with many other event datasets that have been designed to account for news coverage bias. An analysis comparing GDELT's report of violent events in Syria with Syria Tracker's, a crowdsourced project to track violence in Syria, finds a correlation of .53 between the two [Masad, 2013].[14] Similar comparisons do not exist for ICEWS, but ICEWS' correlation with GDELT suggests the results would be the same.

Moreover, even if machine-coded events data record duplicate events, no prob-

[13]This relationship is less strong when the analysis is disaggregated to the city level because ICEWS, and machine-coded events data more generally, is imprecise on subnational geolocation.

[14]The correlation varies depending on which governate one analyses. Geocoding events from news reports is even harder than coding events from news reports, so that the correlations are less reliable than nationwide aggregates is not surprising. The difficulty of geocoding is the primary reason this paper kept analysis at the country-day level.

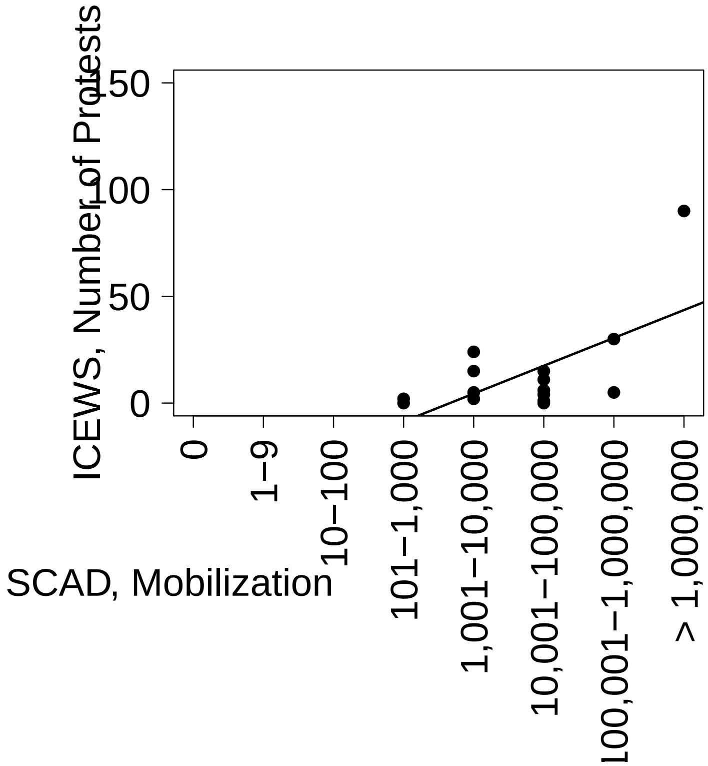


Figure 1.16: More Protestors Correlate with More GDELT Protests

lem exists so long as the duplication is content throughout the dataset. If the noise is a scalar, then the problem is only one of effect size but not direction. For example, if ICEWS multiplies the true number of protests by 5 each time a protest is observed, than the effect found for any variable will be times larger than it actually is, but its direction and relative magnitude will not matter. If the noise of ICEWS is therefore uniform at each moment of time and space, then the noise is a nuisance but does not bias inference.

The noise can threaten inference if it does not vary with a variable of interest or is biased. If it is does not vary with a variable of interest, such as protest, then it can hide any effect of protest that does actually exist. While this outcome would be unfortunate, it increases the risk of a Type II error (false negative). False negatives are unfortunate, but a null finding is not what is being tested, meaning such an effect biases against any finding.[15]

The second concern, the most serious one, is that machine-coded events data may simply replicate bias inherent in news coverage. ICEWS probably over-reports events when an event is new or under-reports an event if it has lasted for a long time. Reporting spikes when an event is new but then tires of reporting on that event. The event therefore will seem to be more widespread than it actually is when it is new but less widespread than it actually is when it has existed for awhile. This problem is especially acute in conflict reporting: what bleeds leads, but continued bleeding is still considered boring. This bias has most recently been shown to apply in the context of the Syrian civil war [Masad, 2013].

Before proceeding, two caveats should be noted. First, the overall effect of news coverage bias is not clear. While it may lead to spurious positive findings at the start of protests, one is less likely to find a correlation, and may find a negative one, if the bias underreports subsequent protests. That is, after n months of protests, individuals still use Twitter at the same levels as before and for the same purpose, meaning any change in Twitter measures reflect true changes in the variable of interest. But these true changes in coordination are then compared to changes in protest counts that may undercount protests if newspapers fatigue of reporting protests. Second, the negative bias in machine-coded data will also exist in handcoded datasets that also rely on newspaper coverage. If a newspaper does not cover an event, then no amount of scholarly intervention will be able to insert that protest into a dataset. Any bias

[15] It creates the risk of a false positive if some other noise component drives the reported number of protests but is not actually driven by the real world number of protests. Whether that is the case or not cannot be known.

that ICEWS exhibits is therefore a reflection of bias in news reporting, but whether the net effect of the two biases cancel out is an open question. Because this study covers time before, during, and after the protest events, the net bias effect could be positive, negative, or neutral.

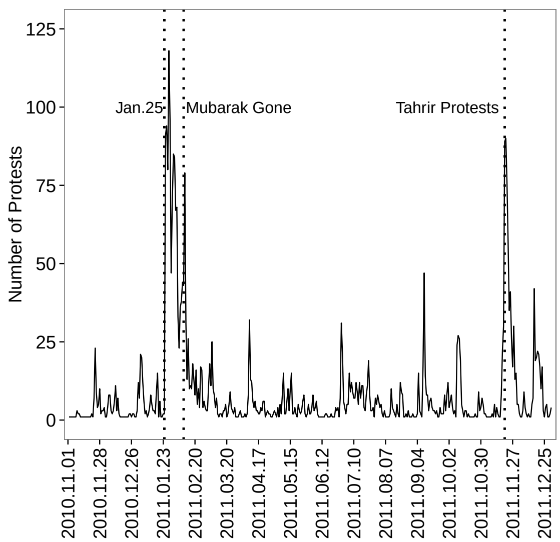
While there is no ground truth against which to compare ICEWS, analysis of the protest patterns revealed in ICEWS suggest news coverage bias does not drive results.[16] ICEWS clearly records protests when they are known to have happened, and it records them at magnitudes which vary according to commonly held beliefs on how widespread those protests were. For example, ICEWS records many more protests for Egypt than it does for Qatar, and it records more protests on January 25^{th}, the first day of protests, January 28^{th}, and the day Mubarak left power than in the summer, though the summer still has higher levels of protest than before January 24^{th}, 2011. Figure 1.17 shows this behavior in Egypt and Bahrain, high-protest countries, as well as Morocco and Qatar, low-protest ones.

Figure 1.17 shows that counts of protest vary and vary in ways that accord with subjective understandings of protests in Egypt and Bahrain. Figures 1.17a and 1.17b show total number of protests for Egypt and Bahrain, respectively. They show a marked increase in protests on January 25^{th} and February 14^{th}, with protest continuing throughout the year. Both countries experience sustained protest throughout 2011 and had their protest movements start on January 25^{th} (Egypt) and February 14^{th} (Bahrain).

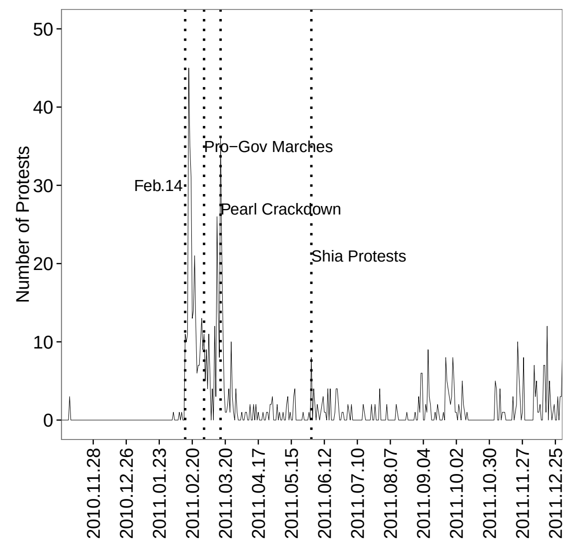
Figures 1.17c and 1.17d look at two countries understood to have experienced little (Morocco) to no (Qatar) disruption during the Arab Spring. Right away, it is

[16]Indeed, ICEWS is the gold standard in machine-coded events data and the closest such dataset scholars have to a ground truth.

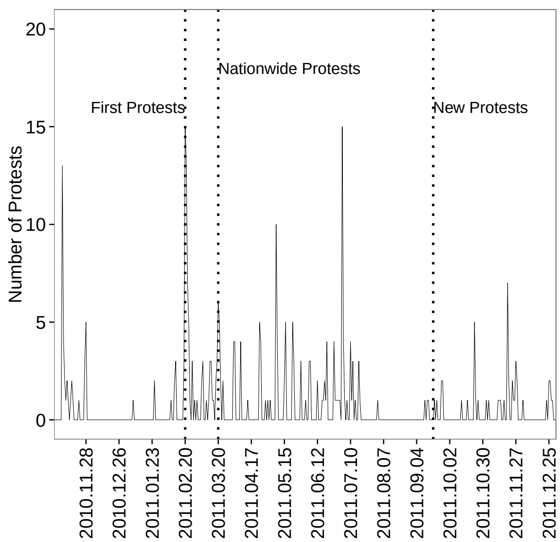
clear that Morocco and Qatar experienced fewer absolute protests, which is expected since they have fewer inhabitants. Both countries seem to have random variance in the number of protests recorded, with Morocco showing distinctive increases on February 20^{th}, the first day of organized protests there, and one month later. Qatar's most distinctive increase occurs on January 25^{th}; as no other sources report protests that day in Qatar, that increase is most likely a spurious correlation with Egypt's protests.



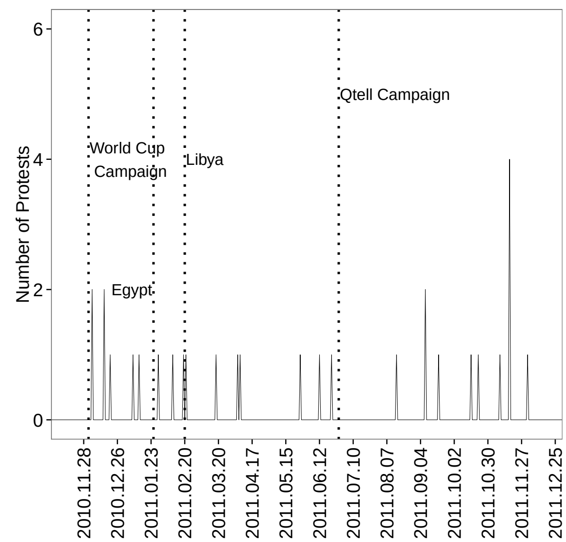
(a) Egypt, Total Protests



(b) Bahrain, Total Protests



(c) Morocco, Total Protests



(d) Qatar, Total Protests

Figure 1.17: ICEWS Protest Recording, 4 Countries

Unexpectedly, Morocco has much higher levels of protest intensity than Egypt or Bahrain. This finding suggests one of two possibilities. First, Morocco may have actually had more protest than is commonly realized. If that is the case, it is interesting and poses no threat to the paper's conclusions. Second, it may be that Morocco receives little attention from news sources, so the high levels of protest intensity are because it is reported on most often when protests occur. If that is the case, the recorded numbers of protest may have higher variance than they should, as protests that occur but do not attract attention are not reported while those that do attract coverage are covered extensively.

Bias in news reporting is also controlled for with two models. First, the main model is recreated without any day with an inordinate amount of protests reported, as these are days likely to suffer from novelty bias. This model is shown in Table 1.12, and the results do not differ from when the days are included. Second, protests from ACLED are used as the dependent variable in almost the same model. The two differences are that ACLED does not measure state repression and a zero-inflated negative binomial is used under the assumption that ACLED underreports some protest events.[17] These results are reported in Table 1.14. $Coordination\_i, t-1$ is still positive but no longer statistically significant, while core coordination is still not significant. A positive effect is found for the percent of tweets with hashtags that are from the core, while a larger negative effect is found for tweets from the core that mention another user. These results should be weighted very slightly, as ACLED only records 5 countries from the study's main sample.

[17]Zeroes are modeled with only an intercept, though rerunning the model with zeroes as a function of an intercept and country fixed-effects does not change results.

Table 1.14: ACLED Protests

	Protest
Coordination$_{i,t-1}$	0.678
	(1.543)
Hashtag %$_{i,t-1}$	1.031
	(1.924)
Retweet %$_{i,t-1}$	−1.244
	(1.977)
Link %$_{i,t-1}$	0.664
	(0.526)
Mention %$_{i,t-1}$	−0.056
	(0.972)
Protest$_{i,t-1}$	0.291***
	(0.085)
Repression$_{i,t-1}$	0.172***
	(0.096)
Core Hashtag %$_{i,t-1}$	2.071***
	(0.796)
Core Retweet %$_{i,t-1}$	0.054
	(0.534)
Core Link %$_{i,t-1}$	−0.632
	(1.144)
Core Mention %$_{i,t-1}$	−2.076*
	(1.187)
Coordination$_{i,t-1}$*Elite Hashtag %$_{i,t-1}$	4.291
	(3.696)
Intercept	−3.029***
	(0.513)
Country FE	Yes
N	2,069
Log Likelihood	−1,241.746

*p < .1; **p < .05; ***p < .01

1.10.8 Deeper Analysis of Egypt, Bahrain, Morocco, and Qatar

This section compares two high-protest countries, Egypt and Bahrain, to two low-protest ones, Qatar and Morocco, to demonstrate how coordination occurs.

Figure 1.18 shows how the prevalence of common hashtags varied in Bahrain, Egypt, Morocco, and Qatar. This figure shows the percent of a day's tweets which contain at least one of the hashtags indicated in that country's legend. The hashtags were chosen by looking at each hashtag that was the most common hashtag for at least one day in each country; The 4 most common hashtags are shown here, except for "#lulu" in Bahrain (not in that country's top 4) because it was the main hashtag used to coordinate action at the Pearl Roundabout. Individuals in Bahrain and Egypt show higher levels of hashtag use than those in Morocco and Qatar; Egypt and Bahrain also had many more protests (see Figure 1 and Table 4 from the main paper).

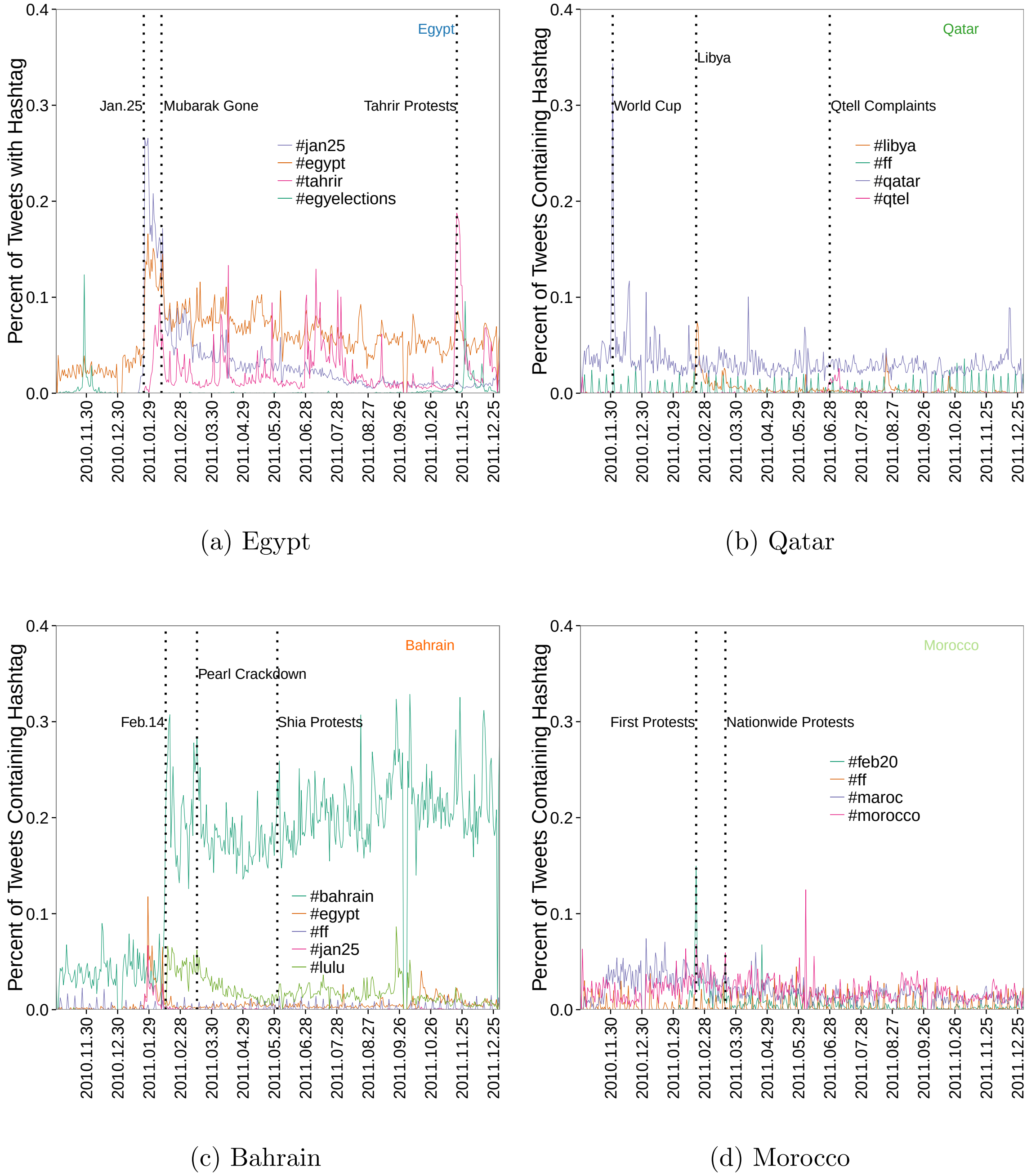
In Egypt, three behaviors stand out. First, little hashtag use occurs before January 25^{th}. "#egypt" is barely used, the first appearance of"#jan25" is not until January 19^{th} (and then is only .1% of tweets), and "#tahrir" does not appear until January 25^{th} (but not in large numbers until just before the resignation of Mubarak). Second, which hashtag is most prevalent depends on the type of upcoming event. During the 18 days of initial protest, "#jan25" dominates, as this was the focal date for the events. Though the largest protests while Mubarak was in power took place in Tahrir Square, that time was a nationwide contentious event, which is why more general hashtags such as "#jan25" and "#egypt" prevail. Moreover, "#jan25" consistently declines in usage after the 18 days of protest. By the middle of March, it will never be the most common of the three hashtags again, and it ceases to correlate with the other two. Third, overall levels of hashtag use decrease after the 18 days. They first decrease sharply after President Mubarak's resignation, and their average prevalence gradually continues to decline. "#tahrir" might be an exception, as it is used much more narrowly to coordinate specific, Tahrir Square-centric events, but the frequency of its usage declines as well.

In Bahrain, there is a dramatic spike in hashtag use almost immediately after the start of protests on February 14^{th}. This hashtag use is then sustained throughout the year, with increases and spikes around important events. "#lulu" is not in the dataset before February 14^{th}, and its use varies over time in more specific patterns than "#bahrain". Finally, note the focus on events in Egypt at the end of January and starting again in mid-October.[18]

Morocco and Qatar experienced less protest. The initial protests in Morocco occurred on February 20^{th}, and only two spikes of hashtag use occur. Moroccans do not appear to have taken interest in events occurring elsewhere in the Middle East and North Africa, suggesting that any grievances they felt were not as salient as they were in countries that did often talk about foreign events. Qatar exhibits similarly low levels of coordination. The day with the highest level of hashtag use is December 2^{nd}, 2010, when Qatar learned it was chosen to host the 2022 World Cup. It is also interesting that events in Libya are the most common topic the 3^{rd} most often, though events in Egypt are 6^{th} and 7^{th}.

The same hashtags correlate with protest in each country. For Figure 1.19, the 4 most common hashtags per country were plotted against the coordination level of

[18] #ff stands for "#FollowFriday". On Twitter, it is common for users, every Friday, to tweet about other users they find particularly interesting and use the "#ff" hashtag to suggest people follow the mentioned user.



(a) Egypt

(b) Qatar

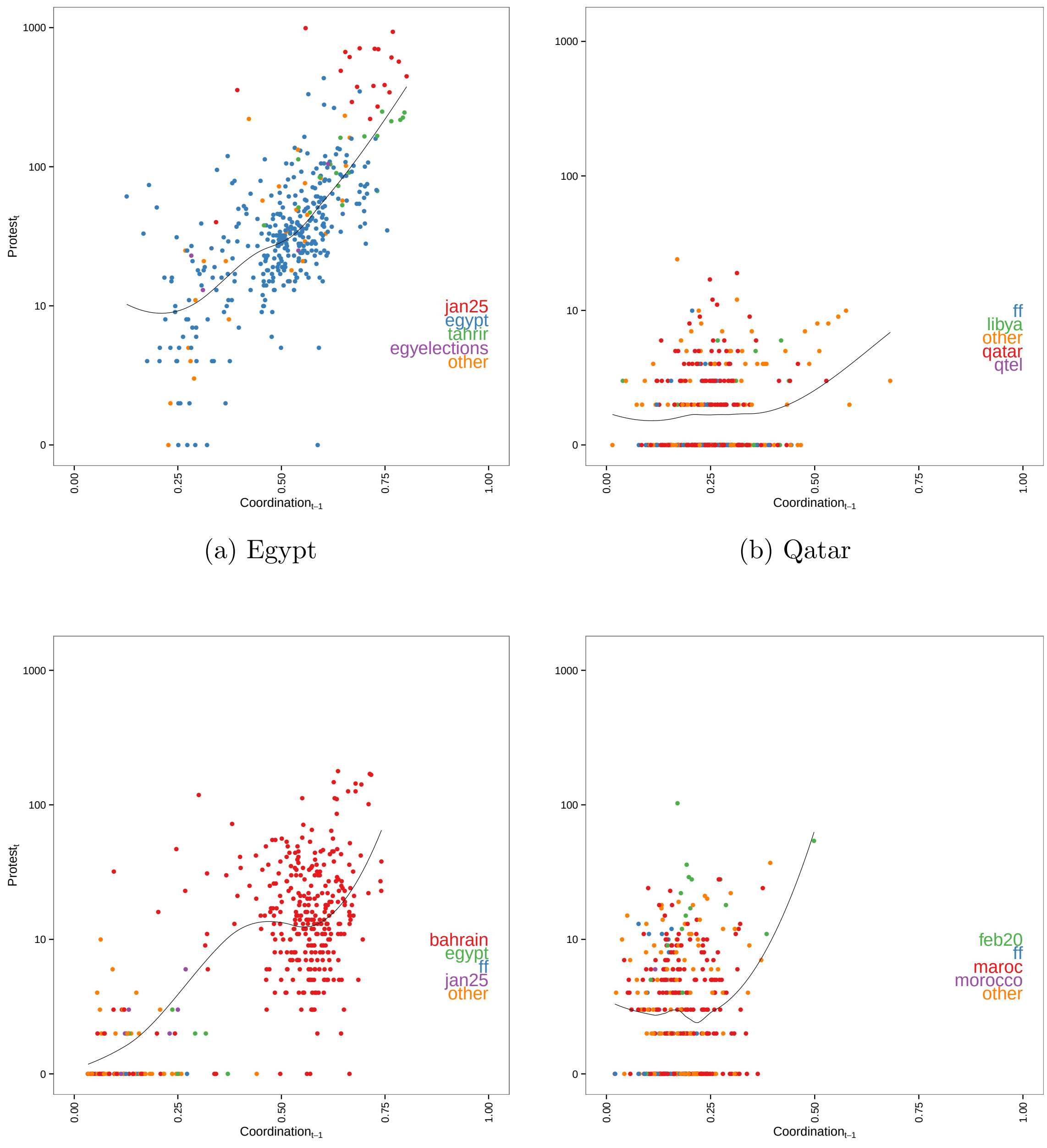
(c) Bahrain

(d) Morocco

Figure 1.18: Hashtag Intensity, Word Varies Based on Events

those hashtags against protest the next day. The figure shows both the differing levels of protest across the countries and the particularistic use of hashtags in each country.

The highest protest days in Egypt and Morocco are associated with protest-specific hashtags, though the most common hashtag for protest and non-protest events in Bahrain is “#Bahrain”.



(a) Egypt

(b) Qatar

(c) Bahrain

(d) Morocco

Figure 1.19: Hashtag Use Varies by Coordination

An examination of the 2 most common hashtags from Egypt and Bahrain and examine the 100 most common words associated with each. English and Arabic tweets were separated, and the Supplementary Materials show how discourse across the hashtags varied.

1.10.9 Shared and Unique Words for Hashtags

This section shows the words shared for each set of hashtags in Egypt and Bahrain as well as the words that were used only in each of the hashtags. Results are shown for English and Arabic sets of tweets.

- **Egypt**
  - *Words Common to #egypt, #jan25, and #tahrir*
    - Arabic (52)

      #egypt, #jan25, #tahrir, on, #25jan, egypt, #noscaf, #news, mubarak, after, ..., #scaf, revolution, tahrir, council, to, today, people, neither, allah, army, between, was, one, egyptian, president, minister, slave, what, other than, mohammed, because, via, #mubarak, police, square, before, now, #ikhwan, #salafi, honour, when, #jul8, the protesters, why, security, youth, until, #ontveg, #may27
    - English (62)

      #egypt, #jan25, #tahrir, time, will, egypt, now, egyptian, #cairo, protest, peopl, say, just, new, #scaf, one, can, get, via, revolut, #mubarak, like, need, day, mubarak, armi, militari, today, know, ..., make, #25jan, want, polic, think, good, pleas, check, love, call, live, see, still, #noscaf, tahrir, come, take, back, right, video, use, watch, start, support, happen, ppl, hope, squar, look, give, help, attack
  - *Words Unique to #egypt*
    - Arabic (18)

      someone, #egynews, #ahram, i am, direct, network, #zamalek, #enn, #7urreya, in it, #6april, #syria, #cairo,#ahly, #libya, like that, #egyfootball, prayer

    * English (24)

      2011, current, city, pray, prayer, #egyptian, service, may, #libya, #egyelect, #egypt., elect, news, beta, country, #egypt:, wednesday, arab, sunday,, #israel, saturday,, follow, tuesday, year

  - *Words Unique to #jan25*
    * Arabic (11)

      #amndawla, january, there are, seventh, #dostor2011, arabian, he said, constitution, shafiq, ahmed, through

    * English (17)

      secure, sinc, regime, preside, talk, @ghonim, #amndawla, force, twitter, must, thug, said, freedom, thing, govern, #egyarmi, monist

  - *Words Unique to #tahrir*
    * Arabic (20)

      #tahrirtv, mahmoud, #nov18, we, #occupycabinet, inside, #tahrir, #alex, rebels, with, #fuckscaf, live, field marshal, #sep9, tear gas, who, #suez, street, down with, on him

    * English (27)

      #tahrir., gas, #jul8, tear, #tahrir,, chant, moham, #may27, mahmoud, number, supply, join, anyone, guy, around, plz, #occupycabinet, #july8, scar, tell, mani, now., fire, #tahrirne, got, tweet, show

- **Bahrain**
  - *Words Common to #bahrain, #feb14, and #lulu*
    * Arabic (48)

      #bahrain, on, #14feb, bahrain, #feb14, #lulu, allah, #gcc, #kuwait, #ksa, today, to, #alwefaq, this, neither, people, before, #egypt, regime, agreement, sheikh, now, hamad, day, was, #bh, salman, until, mohammed, #uae, between, martyr, troops, channel, march, home, report, other than, picture, #q8, village, youth, bahraini, video, jacket, eisa, rajab, noble

    * English (48)

      protest, will, police, now, people, can, attack, ppl, just, right, get, one, call, want, like, see, say, day, know, today, use, arrest, love, need, video, via, think, make, new, back, support, live, still, come, stop, time, good, pleas, tri, ..., happen, start, take, head, watch, thank, look, hope

  - *Words Unique to #bahrain*
    * Arabic (27)

after, @alfarooo8, king, security, but, iran, basyuni, by, world, minister, president, only, #syria, men, all, one, @14febtv, because, someone, ministry, and from, not, and allah, interior, then, #jun1, inside it

  - English (16)

    make, good, #bahraindeb, country, ask, world, year, work, gas, never, must, king, look, even, old, hope

- *Words Unique to #feb14*
  - Arabic (14)

    mercenary, #news, #ajstream, #world, #bnn, #un..., #feb14..., near, #june1, cinder, hasan, street, high, square

  - English (21)

    #arabspr, #bhn, @nabeelrajab, #eu, #bnn, secure, hospit, ali, death, crackdown, tortur, demonstr, tear, die, anoth, student, activist, medic, shot, brutal, shoot

- *Words Unique to #lulu*
  - Arabic (13)

    #arab, village, #lulureturn, that, #usa, #manama, ..., #wefaq, #btv, #lulu..., pearl roundabout, #saudi

  - English (16)

    #lulureturn, roundabout, #freedom, thug, car, #arab, #ksa, tri, #kuwait, demand, @emoodz:, confirm, #bh, #lulu,, pearl, thank

1.10.10 Topic Model Detail

The primary problem with relying on hashtags is that the researcher has to surmise meaning from the tag, and the tag can be attached to texts with wide ranging meaning; only for the most specific hashtags, such as #postegyptianrevolutionsocialtrends can one safely assume a strong correlation between the hashtag and the meaning of its content. To more precisely measure meaning, one has to create a topic model. A topic model is a statistical algorithm that determines how features of a document words, sets of words, syntax, etcetera correspond to the topic of the document.[19]

To create the supervised topic models, 3,000 tweets from Egypt and Bahrain (6,000 total) were handcoded. The coding for Egypt was performed by a team of 3 undergraduates, all native Arabic speakers, who were asked to assign a tweet to any of 40 categories; a tweet could belong to multiple categories. The coders agreed on 95.09% of the tweet-categories. Of the categories directly related to protests, agreement ranged from 66.05% (if a tweet was political but not about protests) to 97.35% (if a tweet was about economic security. Intercoder reliability was 84.00% for tweets about protest coordination and 90.51% for tweets providing common knowledge. Because of financial constraints, the coding for Bahrain was performed by the author and a colleague. The author coded 1,500 English tweets, the colleague, who conducts fieldwork in the Middle East, 1,500 Arabic ones. See the Supplementary Materials for an explanation of these categories.

Upon completion of coding, a supervised topic model was created. A supervised approach was chosen, for 3 reasons. First, the level of interpretation required of

[19] "Document" means the textual unit of analysis. In this study, the document is the tweet, but it can be any text: a speech, a magazine article, a collection of articles, a Facebook post &c.

unsupervised approaches leaves results of those models difficult to interpret. In the unsupervised approach, one takes a collection of documents, tells the computer to how many categories the documents belong, and the computer sorts the documents into those categories depending on a loss-minimization criterion. The number of categories is arbitrary, and the researcher has to test different numbers to find which appears to best divide the documents into natural categories. The researcher then has to interpret the sorting of the documents to understand what real-world topic the groupings represent. Second, even if the interpretation of each category is not contentious, the number of categories is. There is no clear rule to distinguish between 5, 10, or 100 categories. While the supervised approach also relies on choosing a number of topics to which each document could belong, that decision is driven by theory, the researcher's contextual knowledge, and an iterative reading of the documents. Third, supervised learning allows the researcher to define the categories in which one is interested. Unsupervised approaches require the researcher to fit a collection of documents to a grouping that may correspond to a real-world category, whereas the supervised approach fits real-world categories to a document that may or may not actually be about that topic. The latter is therefore best when one knows for what one is looking, such as tweets coordinating protest or talking about a state's repressive response. See Grimmer (2013) and Lucas (2015) for more detail on text analysis.

After deciding upon a supervised topic model, a grid search was performed to tune model parameters. Specifically, a support vector machine and 3 varieties of Naive Bayes classifiers were tested. Each were tried with varying numbers of document features, with the features defined as 1, 2, or 3 n-grams. Those combinations were then trained on 95% of the coded tweets and tested on the remaining 5; the specific

95% was chosen at random with replacement. The training was repeated a variable number of times, from once to fifty, and a tweet is classified as belonging to a category if more than half of the models assign it to that category. This process is known in machine learning as a bagged ensemble, is equivalent to bootstrapping, and lowers the variance of a model. This entire process was performed 4 times: once on protest coordination tweets in Egypt and Bahrain, and then again for common knowledge ones in each country. The model with the highest F1 score was chosen and used to generate predictions for the out of sample data. The final classification parameters are shown in Table 1.15.

Table 1.15: Topic Models' Parameters

Country	Topic	Classifier	Features	N-grams	Bags	F1
Bahrain	Coordination	Bernoulli Naïve Bayes	450	3	30	.65
Bahrain	Common Knowledge	Bernoulli Naïve Bayes	600	1	15	.62
Egypt	Coordination	Bernoulli Naïve Bayes	800	2	25	.64
Egypt	Common Knowledge	Bernoulli Naïve Bayes	900	3	10	.59

1.11 Acknowledgements

Chapter 1, in part, has been submitted for publication. I am the primary investigator and author of this paper.

Chapter 2

Activism and Protest in Authoritarian Regimes

2.1 Abstract

This paper develops a theory of activism in authoritarian regimes to explain how structural conditions can lead to protest mobilization. Activists, individuals specializing in policy change through non-governmental mechanisms, have received less attention from scholars than other components of the mobilization process. Mobilization requires common knowledge and coordinating information, both of which activists provide. For reasons elaborated in the theory, activists are effective at generating common knowledge but not coordinating protest events. This theory is tested using a novel combination of Twitter data and case studies. Case studies show how activists were integral to generating common knowledge in Egypt. Complete Twitter activity is obtained for 19 activists from Bahrain and Egypt from January 11^{th}, 2011 through April 5^{th}, 2011 show that activists' coordinating behavior did not affect subsequent protest mobilization. Instead, activists were most effective negotiating policy with government effectors. Protests are a primary source of policy change in authoritarian regimes, and activists are prominent actors involved in them. Understanding how activism works and how it affects protest therefore provides a new way of understanding policy change in authoritarian regimes.

2.2 Introduction

What causes protests in authoritarian regimes? Scholars have focused on both structural [Gurr, 1971, Skocpol, 1979, Lichbach, 1998, McAdam, Tarrow & Tilly, 2001, Beissinger, Jamal & Mazur, 2015, Chenoweth & Ulfelder, 2015] and individualist explanations [Davenport & Trivedi, 2013, Snow & Moss, 2014, Trejo, 2014]. Protests become more likely as certain structural outcomes change [Chenoweth & Ulfelder, 2015], but individualist models treat the start of protests as an exogenous shock [Kuran, 1991, Lohmann, 1994]. This paper argues that scholars have overlooked a key mechanism that translates structural-based mass dissatisfaction into protests: activists.

This paper suggests an answer to this question through a theoretical and empirical innovation. Theoretically, it emphasizes the importance of activists, individuals who specialize in activity they expect to hasten policy change. Protests are more likely to occur the more that individuals know others share their anti-status quo preferences (common knowledge) and the more those individuals can coordinate protest action (coordination). Activists therefore work to generate common knowledge through activities like creating Facebook pages for people to air grievances, holding workshops, or distributing fliers. They also attempt to coordinate protests by suggesting spaces in which to congregate, materiel to bring, and how to interact with state forces. The nature of common knowledge generation and coordination differ in crucial ways, however, and these differences mean that activists mobilize individuals through generating common knowledge, not coordinative activity.

The theory is tested combining high-resolution social media data on 19 activists with case studies of activists' behavior. These activists are from Bahrain and Egypt during the 3 months at the beginning of 2011, a period which encompasses the start of Arab Spring protests and their primary resolution in both countries. To quantify activists' behavior, their entire Twitter history from this period is used. These data are supplemented with primary source material to understand their behavior outside of social media. These qualitative data complement social media data by giving context to the changes in behavior observed via Twitter.

Moreover, many studies of mobilization also focus on low-risk activities, such as signing a petition or donating to a campaign [Brady, Schlozman & Verba, 1999]; activities where individuals risk physical harm, such as protests, follow a different logic [McAdam & Paulsen, 1993, Centola & Macy, 2007]. Even fewer studies analyze this recruitment process in authoritarian regimes [Opp & Gern, 1993, Beissinger, Jamal & Mazur, 2015]. How structural preconditions translate into actual protest in authoritarian regimes therefore remains understudied, though it has received more interest from scholars since the start of the Color Revolutions in the mid-2000s [Tucker, 2007, Mungiu-pippidi & Munteanu, 2009, Bellin, 2012, Steinert-Threlkeld et al., 2015].

This paper proceeds in six main parts. Section 2.3 develops a theory of activists' behavior. Section 2.4 introduces the data sources and measures used to test the theory. Section 2.5 presents qualitative results about common knowledge creation and quantitative ones for activists and coordination. shows the quantitative results and supplements them with primary source material. Section 2.6 explores differential rhetorical patterns of activists and non-activists as well as discusses effects activists may have during protest periods. Section 2.7 summarizes the results and discusses implications of the theory.

2.3 Theory

Activists

Activists are individuals specializing in generating policy change outside of formal organizations such as political parties, unions, or mass media. While "political entrepreneur" is the term commonly used to describe activists, that term is too broad because it encompasses individuals who may work through traditional channels such as a legislature or political party. Activists, on the other hand, try to change policy through generating grassroots support that leads to mass political action.

Though a particular individual may have grievances against the regime or recognize that policies are far away from his or her preferred policy, an individual does not know if those grievances are shared by others or attributable directly to the regime. Once individuals know that others share these grievances, they are more likely to protest because they know others are likely to protest as well. This shared knowledge is known as common knowledge.

Mass media provide an ideal platform for generating common knowledge. The increasing returns to scale of information production from centralized media means that the most efficient way to make individuals aware of shared dissatisfaction is through dissemination via the radio and television [Warren, 2014]. But authoritarian regimes routinely control these platforms, rendering them unavailable for the production of anti-regime common knowledge [Egorov, Guriev & Sonin, 2009]. Common knowledge must therefore be generated via another mechanism.

Activists are the common knowledge mechanism. For example, many Egyptians had negative experiences with the police, but there was a belief that these experiences applied primarily to the poor and working class. The police's public beating of Khaled Saeed, a middle-class Egyptian teenager, in broad daylight, and the subsequent knowledge of the beating through social media, came to represent that everyone was vulnerable to the regime's capriciousness [Cambanis, 2015]. In the United States of America, race relations between African-Americans and the police have long been contentious, but it took the near-simultaneous deaths of Michael Brown and Eric Garner to instigate a nationwide movement and dialog about the issue. Activists work to spread information about such events, generating common knowledge between individuals that many others share their dissatisfaction. Those who thought few others share their anti-status quo policy preferences begin to realize many others do.

Activists are uniquely situated to generate common knowledge because they specialize in information dissemination outside of state control. If a protest is conceived as a kind of global game [Carlsson & van Damme, 1993], then limited depth of reasoning can prevent mobilization from occurring. The role of activists is then to render private information public, generating an equilibrium in which mobilization occurs [Edmond, 2013, Kneeland, 2014]. This argument is similar to why advertising can be effective even if the advertisement contains little information about the product: when the product gains value as more people use it, a product's success is tied to how many people know about it in the first place [Chwe, 1998]. Similarly, "advertising" anti-regime preferences generates common knowledge that many others know about anti-regime preferences. The generation of common knowledge about anti-regime preferences creates a reservoir of individuals primed to mobilize [Klandermans & Oegema, 1987].

These activities can include holding workshops about peaceful resistance, showing movies about protest or critical of a regime, and creating political graffiti are examples of these kinds of behaviors. These activities demonstrate to others that there exists discontent with a regime, and they do so in a way so that others know that others know there is discontent. This common knowledge helps resolve the coordination problem of knowing that others want to, and will, protest in the same public information resolves the coordination problem of speculative currency attacks [Kneeland, 2014]. Generating common knowledge has the same affect as advertising: the more people know about a product (the protest), the more likely are people to buy it (join the protest) [Lohmann, 1994, Chwe, 1998, 2003]. For example, Egyptian activists deliberately discussed (advertised) the January 25^{th} protests in taxi cabs because they knew the drivers would spread the information to their passengers [Lim, 2012].

Once a mobilization potential is established [Klandermans & Oegema, 1987], potential protesters need to become actual protesters. Protest occurs once everyone can coordinate the details of a particular protest, such as when to have it, where to go, and what to bring. Even when elections and national holidays provide this coordination , individuals must still decide other particulars [Fearon, 2011]. For example, Egypt's first protests were scheduled for National Police Day, January 25^{th}, but that did not determine the time of day or tactics used. In addition, authoritarian regimes forbid or co-opt organizations that could mobilize masses against it. Opposition parties can be used to stifle movements by buying off leaders [Reuter & Robertson, 2015], and labor unions are similarly used to monitor and control potential threats [Robertson, 2007, Bishara, 2014, Chomiak, 2014]. Coordination must therefore be generated via another mechanism.

Much literature examines how social movement coordination is created, and a common finding is that activists in the movement are key mobilizers. In addition to creating common knowledge, they appear to also tell people to engage in specific action, either immediately or in the near future. For example, getting a non-mobilized individual to talk about a political action can cause that individual to take that action [Klofstad, 2007]. Activists also engage in recruitment, asking individuals to partake in a civic action [Abramson & Clagget, 2001], and they strategically target susceptible individuals [Brady, Schlozman & Verba, 1999]. This recruitment often occurs along social network lines [Snow, Zurcher Jr. & Ekland-Olson, 1980, **?**], though there is some debate about whether these connections have an effect independent of other factors [McAdam & Paulsen, 1993]. Outside of political action, people who interact many other people (central nodes, in network terms) spread behaviors in a population more than others [Kim & Bearman, 1997, Watts, 2004].

Those findings should not apply everywhere and for all types of recruitment, however; activists should not have a coordinating effect in authoritarian regimes. The empirical literature on recruitment to social movements focuses on low-risk endeavors, like signing a petition or donating to a political party, in democracies [Brady, Schlozman & Verba, 1999]. These kinds of activities are highly unlikely to pose risk to the recruited, making them more likely to partake. High-risk activity, such as a Freedom March or protest, will require more inducement, even in a democracy [**?**Davenport, 2007*c*]. But a protest or sit-in in a democracy is still lower risk than in more repressive regimes [Andrews & Biggs, 2006]. High-risk activity in a regime which represses that activity, like protest in an authoritarian regime, should therefore be even harder to induce. People will not be convinced to join a protest in Tripoli when a stranger approaches them on the street.

In authoritarian countries, protesters face a much greater risk of physical harm than in democracies [Hill Jr. & Jones, 2014]. Even if primed to mobilize through common knowledge, mobilization requires 2 pieces of information: that those mobilizing come from different parts of society, suggesting protests will be large [Lohmann, 1994, Ugander et al., 2012], and, more prosaically, logistical information about where to go, when to go there, and what to do. These 2 pieces of information allow individuals to coordinate about impending and ongoing protest. Only by engaging in coordinated mass action will the non-activist mobilize, unlike in a democracy. These two pieces of information are better provided by non-activists than activists.

Activists do not have a coordinating effect because they are too few in number. Being few in number, their mobilization provides little to no evidence that other kinds of people, like factory workers or the unemployed, are also mobilizing. On the other hand, a large protest is seen as having many people who do not usually protest, so non-participants judge that the protest has a higher chance of success than expected, and people who would not normally join are more likely to participate [Lohmann, 1994]. In terms of logistical information, their paucity means the information they provide about places to protest, routes to take, and what to bring will be less accurate than that from the the other participants also providing the same information. This logic is explored in more detail, and tested, elsewhere [Steinert-Threlkeld, 2016].

The importance of non-activists for coordination has been explored extensively in the context of East Germany's 1989 protests. There, mobilization appeared to self-catalyze, as the likelihood of participation increases as the likelihood of success is believed to increase as well [Finkel, Muller & Opp, 1989]. In Leipzig, "incentives to participate were concentrated in networks of friends" since "communication is easy, and members expect approval for participation in protests, particularly if members have close ties" [Opp & Gern, 1993, pgs. 659, 662] Indeed, the authors also point out that "the protests were not a result of the mobilization activities of opposition groups" (page 671). People mobilized when those like them mobilized because they shared the same identity and grievances, and the institutional structure in East Germany made it difficult to learn about these grievances outside of ones immediate social network. Moreover, activists were few in number, actively repressed, and the civil society they would build derived from - it did not lead - the mass mobilization [Pfaff, 1996].

The importance of activists for common knowledge creation but their weakness at coordination leads to the following hypotheses:

H1: As activists provide more information about anti-status quo preferences, more protests will occur.

H2: As activists provide more information about protests, no more protests should be observed.

Why Authoritarian Regimes?

This paper focuses on activists in authoritarian regimes because that is where they are most likely to be the primary generator of common knowledge. In countries with high levels of media freedom and where individuals can gather freely, activists compete with other groups to generate common knowledge. At the other extreme, in totalitarian countries, a dearth of civil society organizations and widespread surveillance prevent activists from generating common knowledge. In between are authoritarian regimes, countries where the state controls most, but not all, political, economic, and media institutions. These states also work to marginalize activists, but they permit them some activity under the belief that it prevents more serious political challenges [Gandhi & Lust-Okar, 2009, Miller, 2014, Miguel, Jamal & Tessler, 2015].

In these countries, activists are most likely to affect protests through the generation of common knowledge.

In mature democracies, voting is the most common method of policy change. The aggregation of preferences through defined rules, and the willingness of those in power to heed the result, has many advantages. It is a low-cost endeavor for participants, as the only costs are transaction and opportunity costs. Little coordination is required, as voting places and times, as well as the choice set, are well-defined. The ability to peacefully remove leadership without changing institutions promotes economic and social stability. On the other hand, an individual vote has an infinitesimal impact on the final outcome [Downs, 1957], usually does not change the structure of a political system, and may change policy more slowly than other alternatives.

In authoritarian regimes, however, voting is a less significant act. In countries where policy feedback comes from an insider population drawn from the larger populace [Bueno de Mesquita et al., 2003], those belonging to the outsider population provide policy feedback through rebellion or protest. Moreover, individuals in democracies face much less repression than in autocracies. They can sign a petition or join an advocacy organization [Klandermans & Oegema, 1987]. They gather in groups with little fear of arrest [Kim & Bearman, 1997]. Political challengers can put forth policy platforms that differ markedly from the prevailing status quo; they can even ascend peacefully to power and implement those policies [Pappas, 2008]. These actions are not practical in authoritarian regimes.

While protests occur in democracies as well, they occur less frequently because fewer groups are excluded from policy making than in autocracies. Because of this difference, mobilization in autocracies differs from in democracies: in the former, the goal is to mobilize people from different sectors of society [Ugander et al., 2012], while in democracies the goal is to mobilize those already inside the political system. In the context of recruitment to political causes in the United States of America, Henry Brady, Kay Schlozman, and Sidney Verba remark that:

> political recruitment does not usually mobilize excluded constituencies to politics. Rather, the overall thrust is to reinforce the tendencies of a participatory process anchored in individuals' political engagements and resources. [...] The process of citizen recruitment brings into politics activists who closely resemble those who would have taken part spontaneously. Political recruitment does not mobilize the marginal and disposed. [Brady, Schlozman & Verba, 1999, pg. 154]

Later in that same article, they summarize:

> When viewed in toto, however, our data suggest that processes whereby citizens are recruited are less likely to transform than to replicate the representational outcomes of a participatory system in which the privileged speak more loudly than the disadvantaged. Indeed, by recruiting activists on the axis of the same factors that would lead individuals to take part on their own, the recruiters are likely to amplify the effects of those factors on participation. (pg. 165)

The statements from Brady et. al. do not apply in authoritarian regimes because most individuals there are excluded from politics. Verba et. al, and the larger literature on participation in democracies, assumes that barriers to entry are low. In an authoritarian regime, who participates in political decision making is tightly controlled, and participation is not allowed without an invitation. There are no individuals who would would have participated without recruitment because there is no participation. Participation has to be created, and activists generate it through a long-term process of common knowledge creation.

To summarize, authoritarian regimes restrict contestation through the ballot box, so political dissent finds other ways to express itself. Activists are the mechanism by which this dissent becomes expressed, and the political opportunities available give them advantages not present in democracies or authoritarian regimes.

Why Protest?

In authoritarian regimes, protests are the most likely expression of political dissent because the cost of rebellion is too high for most individuals and voting is ineffective.

Almost every human lives in a large group that requires an aggregation mechanism to translate preferences over events into action that governs that group.[1] The aggregation mechanism, resulting outcome, and subsequent contests over each is called politics. In any system in which aggregation occurs, the resulting policy will necessarily differ from many individuals' desired policy. Assuming that each person wants a policy that exactly matches his or her preference, contestation ensues to change the policy.

There are three ideal types of contestation: voting, mass violence (rebellion), and mass non-violence (protest). The demand for these forms of contestation increases as the distance between the median citizen's policy and the prevailing policy increases. Individuals have different thresholds beyond which they benefit more from contentious action than inaction, though these thresholds are not directly observable, their distribution varies in a population, and this distribution affects the likelihood of contestation's start [Granovetter, 1978, Centola & Macy, 2007]. Precisely when

[1] The group depends on the level of analysis. Any two people working together, whether as roommates, colleagues, teammates, &c, are engaged in this aggregating activity and so form a group. The most common group is the family, and preference aggregation is simple. The extent of politics then increases with the size of the group, with the state being the largest sovereign political group.

protest or rebellion starts is a frequently studied question, with scholars focusing on both structural [Gurr, 1971, Skocpol, 1979, Lichbach, 1998, McAdam, Tarrow & Tilly, 2001, Fearon & Laitin, 2003, Chenoweth & Ulfelder, 2015] and individualist explanations [Davenport & Trivedi, 2013, Snow & Moss, 2014, Trejo, 2014]. After starting, states routinely respond with overt repression [Rasler, 1996, Moore, 1998, Davenport, 2007*a*, Boudreau, 2009], even in democracies [Davenport, 2007*c*].

A rebellion is an armed, violent movement against institutions of the state and does not have to be revolutionary. A rebellion provides a state with a strong signal of discontent and can be costly to the regime, both it terms of foregone economic activity and physical damage to state capabilities. A rebellion is difficult to scale, however, as it has high financial cost and puts participants in mortal danger; it also requires material and financial support from the population in the area it operates, which in turn requires high levels of consensus.[2]

A protest is a gathering of people expressing desire for policy change. Protests can send strong signals about policy preferences to a state if enough participants join the protest, as that implies that participants come from different sectors of society [Lohmann, 1994]. Repressing a protest, especially one that is large and peaceful, then imposes high costs on a regime, both in terms of internal and international support; short-term stability is often traded for long-term fragility [Moore, 1998, 2000]. Though protests require some coordination, they are lower-cost endeavors than rebellion or forming a political party, as they do not require the finances of either or programmatic policies of the latter. Protests can therefore mobilize large numbers of disparate individuals at relatively low cost while imposing large costs on

[2]A rebellion based on control of a valuable natural resource, such as coca plants or diamonds, does not fit this schema.

a regime in the form of policy concessions or repression.

Mass mobilization increases the net expected benefit for each participant by lowering the cost of protesting and increasing the likelihood that a policy change (the benefit) is realized [Tullock, 1971]. The cost of protest is lowered because there are returns to scale of protesting and a lower chance of being affected by state repression. Like any undertaking, there are fixed costs to starting and maintaining a protest. While the costs of starting a protest are born by protest organizers, the costs of maintaining it are shared by all participants. Costs include the provision of supplies (food, medicine, and housing), maintenance and guarding of the protest site, and recruitment of additional participants.

Moreover, larger protests decrease the chance a particular individual will experience state repression because it becomes more difficult for a state to repress as protests grow and, if a state does engage in repression, the chance of being arrested or injured is lower in a large crowd than a small one. Susanne Lohmann explains this dynamic when discussing the mass protests in East Germany in 1953 and 1994:

> In the failed 1953 revolt, a few dozen East Germans out of 500,000 strikers and 408,000 demonstrators died. To my knowledge, there were no reported deaths caused by police or security forces for the 1989 East German revolution. Of course, participants in the critical demonstration of October 9, 1989, faced the very real possibility that their protest would end in a massacre; but even then they could reasonably expect that only a few dozen - the wirst case perhaps several hundreds - out of tends of thousands of participants would meet their death. For any one demonstrator, the implied probability of death is rather low. [Lohmann, 1994, pg. 90]

Repression increases the cost of protest, with unpredictable effects on subsequent protest mobilization [Lichbach, 1987, Rasler, 1996, Francisco, 2004, Pierskalla, 2010, Siegel, 2011]. The only regularity in state repression is that it occurs [Davenport, 2007*b*]. With enough repression, however, a state can maintain power in the face of protest, either by using such force that protesters disappear (China in 1989, Bahrain in 2011, Egypt in 2013) or using enough force to hold onto power while protests continue (Syria 2011).[3] The greater the number of people engaged in a protest, the lower these costs are per person.

No amount of mobilization, however, can overcome a determined state. Indeed, a critical turning point in many protest events is when it becomes clear that a regime will not engage in mass violence; realizing the regime, or members of the regime, has a limit to its violence encourages people to protest because the cost of protesting is now much more certain. In many cases, those in power are more than willing to kill as many of those who do not want them in power as is necessary. Burma's mass unrest of 1987 and 1988 led to the resignation of the entrenched Prime Minister, Ne Win; a military coup in September of 1988 led to the death of thousands of students and monks, and the pro-democracy movement was silenced. The People's Republic of China demonstrated similar resolve in the face of mass protests in Tiananmen Square; deploying the army and killing hundreds of protesters signaled that the Communist Party would remain in power. In 2013, mass protests from the supporters of deposed Egyptian President Mohamed Morsi did not stop until the police massacred almost 1,000 participants; in 2011, President Mubarak used force against protesters, but never on the scale General Abdel el-Sisi used that August. A protest will never succeed against a resolute state.

[3] Exactly what "enough" means is an open question, and the nature of state repression is outside the scope of this theory.

An Example

The experience of the Egyptian Facebook page “We Are All Khaled Said” illustrates how these variables operate and the predictions of this theory. Khaled Mohamed Said was a middle-class teenager beaten to death in public by two police officers on June 6^{th}, 2010. The government claimed he died swallowing a bag of marijuana, but a leaked photo of his severely disfigured face (and the presence of eyewitnesses) made it clear he was beaten to death; the police officers faced no immediate repercussions. The page, started on June 10^{th}, 2010 by Wael Ghonim and AbdelRahman Mansour, quickly gathered hundreds of thousands of followers. Its initial purpose was simply to express rage at police impunity; 36,000 people joined on the first day, and one of the first events advertised was Said’s funeral [Ghonim, 2012, pgs. 59-63].

Through using vernacular Arabic, eschewing explicitly political language and affiliations, and focusing on online demonstrations of anger, the page encouraged non-activist Egyptians to voice their political frustrations in a novel, public manner. Less than two weeks later, the page organized silent stands in Alexandria and Cairo; dressed in black and arrayed in a line, participants attracted attention without protesting, and the stands would continue periodically through August. The page quickly took a life of its own, with members posting photos and thoughts on their experiences at the silent stands.

As the year progressed, the page became a hub of political activity, with over 250,000 members by September; Ghonim and Mansour increased their networking with established activists. Yet, at the end of 2010, here were no specific plans for any protests: the page had revealed there were large reservoirs of negative political sentiment, and the publicness of these individual revelations generated common knowledge amongst large groups of individuals that they were not alone in their political beliefs. While January 25^{th}, 2011 was identified as a day to protest, it was not until President Ben Ali of Tunisia fled power on January 14^{th} that the protests became revolutionary. Leading up to the first protest, Ghonim and Mansour used the page to discuss protest tactics, encourage members to recruit those not on Facebook, and presented a list of broad grievances with wide appeals [Ghonim, 2012, pgs. 130-137]. By January 24^{th}, knowledge of the impending protests was widespread, as government officials downplayed the possibility and government Facebook accounts joined "We Are All Khaled Said" to provide anti-protest information there.

In terms of this paper's theory, the "We Are All Khaled Said" page was instrumental in generating common knowledge that the regime of President Mubarak was widely disliked, breaking the "wall of fear" that had prevented Egyptians from contemplating collective action. The page was created by a politically-minded professional (Ghonim) and experienced activist (Mansour), and it provided a platform to generate common knowledge. As Section 2.5 will show, it was not an instrumental tool for coordinating protests, and the primary role of Ghonim, and activists more generally, was not to coordinate protests; once protests started, they engaged in policy negotiation with regime officials.

2.4 Research Design

Selecting Cases

The activism theory is tested using data on 19 activists in Egypt and Bahrain. These countries were selected for two reasons. First, they both experienced high levels of protest in 2011, so periods of protest preparation, protest engagement, and post-protest action should be clearly observable. Second, many activists adopted Twitter before protests started. Four movements were then identified for further study: the April 6^{th} youth (Egypt), No Military Trials (Egypt), anti-sexual harassment (Egypt), and human rights (Bahrain); see Table 2.1 for detail on these actors.

The April 6th movement started in early 2008 as a Facebook page rallying support for striking textile workers at a government enterprise in Mahalla al-Kubra, a city of 535,000 70 miles north of Cairo. Large-scale strikes and protests focused on working conditions and pay had occurred since 2006, sparking a periodic series of worker actions throughout Egypt over the next two years [Beinin, 2009]; a large strike was called for April 6^{th}, 2008, and the government preemptively arrested activists and closed off public spaces nationwide [Gunning & Baron, 2013, pgs. 59-61]. The movement persisted at a subdued level of activity - not for lack of trying - for the next three years and would become a central actor in the 2011 mobilization.

Civilian trials in military courts have been a feature of Egyptian politics since independence [Reza, 2007]. The military court system in Egypt did, and does, not provide civilians with due process; allows for indefinite detention; and often tries defendants in groups within brief trials [Albrecht, 2005]. While military trials of civilians were a tool of the Mubarak regime, their use grew in the weeks and months following Egypt's January 25^{th} protests. The No Military Trials campaign, often called #NoMilTrials because of the movement's preferred Twitter hashtag, gained momentum at this time.

Egyptian public spaces have long been dangerous for women [Amar, 2011]. As protests increased in Egypt throughout the first decade of the new millennium, so did reports of sexual assault at these events; in many cases, these assaults are linked to civilians the Interior Ministry hired for that purpose [Langohr, 2013]. Public celebrations during Eid al Fitr became common sites of sexual assault, leading to a coordinated advocacy campaign for the passage of an anti-sexual harassment law (it would not pass in 2010, though a law making sexual assault a crime was passed on June 11^{th}, 2014) [Ebaid, 2013]. The anti-sexual harassment movement constitutes an assortment of civil society organizations, most notably the Nadeem Center, Egyptian Center for Women's Rights, and the Nazra for Feminist Studies. In December 2010, a new NGO, HarassMap, was created that maps where assaults occur, sends volunteers to those locations to enlist the help of local business; and provides resources to victims of assault. After a brief lull during the 18 day occupation of Tahrir Square, sexual harassment again arose with the security vacuum created by Hosni Mubarak's flight.

Bahrain has had an active human rights community since before the 2011 protests, with the two main organizations, the Bahrain Center for Human Rights (BCHR) and Bahrain Human Rights Society (BHRS), having been founded in 2002. The BCHR is run by Nabeel Rajab when he is not in prison and by family members when he is. A large part of Bahrain's human rights movement, individuals and organizations, is abroad to avoid domestic repression.

News reports, scholarly articles, and NGO documents identified movement activists from Egypt and Bahrain who also had Twitter accounts [Human Rights Watch, 2011, Jones & Shehabi, 2012, Radsch, 2012, Rizzo, Price & Meyer, 2012, Gunning & Baron, 2013]. Because those publications were all in English, it is possible that the accounts will use English more frequently than the average Egyptian or Bahraini. Awareness of Twitter was also quite low in these countries before the start of the Arab Spring, so users tended to be educated individuals, people more likely to be comfortable with English. 42 Twitter accounts were identified, though not all of them were active in early 2011.

Data

Data on activists' online activity were obtained from Sifter, a third-party reseller of historic Twitter data. Sifter provided every tweet from these activists from January 11^{th}, 2011 through April 5^{th}, 2011. This time period was selected because it encompasses time before each country's main protest period, each country's main protest period (January 25^{th} - February 11^{th} in Egypt, February 14^{th} - March 17^{th} in Bahrain.), and time after the protests. The end of the main protest period in Egypt is defined as Mubarak's resignation. In Bahrain, protests ended following a major assault on the Pearl Roundabout, the main protest site in Manama; this assault occurred three days after Gulf Cooperation Council forces, led by Saudi Arabia, marched into Bahrain, and the Pearl Roundabout was dismantled on March 18^{th}. Protesters would not again succeed in occupying the circle. Sifter returned 58,376 tweets; each includes metadata on the number of followers of the account, number of people the account follows, and a character string describing the device from which the tweet was created. Table 2.1 details these data.

Table 2.1: Descriptive Statistics from Sifter Data

Account	Followers	Friends	Tweets	Twitter.com	HTTPs	iPhone	Android	BlackBerry	Windows	Nokia	Retweets	Mention	Hashtag	Group
Shabab6april	4229	78	833	0.67	0.00	0.00	0.00	0.00	0.00	0.00	0.09	0.16	0.87	April 6
mrmeit	2377	244	16262	0.84	0.00	0.00	0.00	0.00	0.00	0.00	0.16	0.59	0.32	April 6
AsmaaMahfouz	1418	88	201	0.98	0.00	0.00	0.00	0.00	0.00	0.00	0.07	0.57	0.16	April 6
waleedrashed	694	329	3265	0.77	0.00	0.00	0.00	0.22	0.00	0.00	0.36	0.76	0.24	April 6
Seldemerdash	1616	277	11832	0.14	0.00	0.67	0.00	0.00	0.00	0.00	0.19	0.54	0.22	Anti-SH
Anti-SHmap	726	154	234	0.51	0.00	0.11	0.00	0.00	0.00	0.00	0.32	0.43	0.38	Anti-SH
SorayaBahgat	315	267	344	0.11	0.00	0.00	0.00	0.89	0.00	0.00	0.15	0.66	0.42	Anti-SH
MariamKirollos	127	37	564	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.52	0.66	0.60	Anti-SH
Ribeska	15	53	23	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.42	0.33	0.83	Anti-SH
ZeinabSabet	2	7	5	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	Anti-SH
alaa	13109	371	38755	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.28	0.92	0.29	NoMilTrials
Monasosh	8595	251	9479	0.47	0.00	0.00	0.00	0.53	0.00	0.00	0.18	0.68	0.40	NoMilTrials
NABEELRAJAB	7285	433	2985	0.80	0.00	0.16	0.00	0.00	0.00	0.00	0.10	0.22	0.84	Hum. Rights
BahrainRights	6687	574	4641	0.00	0.00	0.63	0.00	0.00	0.00	0.00	0.28	0.68	0.92	Hum. Rights
MARYAMALKHAWAJA	4790	187	1112	0.45	0.00	0.30	0.00	0.25	0.00	0.00	0.43	0.54	0.76	Hum. Rights
angryarabiya	3274	144	2575	0.80	0.00	0.00	0.00	0.13	0.00	0.00	0.10	0.62	0.28	Hum. Rights
SAIDYOUSIF	765	56	1574	0.49	0.00	0.51	0.00	0.00	0.00	0.00	0.34	0.47	0.84	Hum. Rights

NB: Accounts are sorted alphabetically within each movement. Columns from Twitter.com through Nokia refer to the percent of tweets from each source; for example, 84% of @mrmeit's tweets are through Twitter's website. Columns Retweets through Hashtag refer to the percent of each account's tweets which are a retweet, contain a user mention, or contain at least one hashtag.

To compare activists' Twitter behavior in context, it is necessary to have data about their country's Twitter behavior at the same time. The non-activists' data can then normalize the activists' data, increasing the chance that findings do not result from spurious correlation of increased Twitter usage simultaneously with increased protests. These data are obtained from a group of researchers who were downloading Twitter data from 2010-2012 [Mocanu et al., 2013]; working with them, tweets from Bahrain and Egypt were selected, and those tweets are used to baseline activists' behavior.

The Integrated Conflict Early Warning System (ICEWS) provides data on protests [Boschee et al., 2015]. ICEWS is a Department of Defense project, led by Lockheed Martin and Michael Ward, that reads newspapers and extracts events. It represents a substantial modification and extension of Philip Schrodt's Kansas Events Data System (KEDS) and Textual Analysis by Augmented Replacement Instructions (TABARI) [Schrodt, Davis & Weddle, 1994, Gerner et al., 2002]. ICEWS reads thousands of news sources, including non-English ones, and applies a heavily modified version of TABARI, leading to much lower rates of false positives than other machine-coded events data.

Measurement

To measure common knowledge, I rely on qualitative analysis of activists' activity leading up protests. I choose this approach because it is not clear how to measure the accumulation of common knowledge without a detailed reading of how activists behaved during each country's pre-protest period. The results of this case study are presented in Section 2.5.

Two variables measure coordination. First, a Gini coefficient of hashtag concentration is used. A Gini close to 1 means most occurrences of hashtags are concentrated on one or a few hashtags, while a value close to 0 means each hashtag occurs in almost equal amounts. The hashtag Gini therefore measures how focused individuals on Twitter are on a topic, so it is used to measure non-activist coordination. See Steinert-Threlkeld et. al. 2015 for more detail on this measure. This variable is used to measure coordination from non-activists.

The second third measure of coordination is a supervised topic model that reads each tweet and decides whether or not a tweet measures coordination, with aggregation occurring the same way as for the common knowledge model.

To create the supervised topic models, 3,000 tweets from Egypt and Bahrain (6,000 total) were handcoded. The coding for Egypt was performed by a team of 3 undergraduates, all native Arabic speakers, who were asked to assign a tweet to any of 40 categories; a tweet could belong to multiple categories. The coders agreed on 95.09% of the tweet-categories. Of the categories directly related to protests, agreement ranged from 66.05% (if a tweet was political but not about protests) to 97.35% (if a tweet was about economic security. Intercoder reliability was 84.00% for tweets about protest coordination and 90.51% for tweets providing common knowledge. Because of financial constraints, the coding for Bahrain was performed by the author and a colleague. The author coded 1,500 English tweets, the colleague, who conducts fieldwork in the Middle East, 1,500 Arabic ones. See the Supplementary Materials for an explanation of these categories.

Upon completion of coding, a supervised topic model was created. A supervised approach was chosen, for 3 reasons. First, the level of interpretation required of unsupervised approaches leaves results of those models difficult to interpret. In the unsupervised approach, one takes a collection of documents, tells the computer to how many categories the documents belong, and the computer sorts the documents into those categories depending on a loss-minimization criterion. The number of categories is arbitrary, and the researcher has to test different numbers to find which appears to best divide the documents into natural categories. The researcher then has to interpret the sorting of the documents to understand what real-world topic the groupings represent. Second, even if the interpretation of each category is not contentious, the number of categories is. While the supervised approach also relies on choosing a number of topics to which each document could belong, that decision is driven by theory, the researcher's contextual knowledge, and an iterative reading of the documents. Third, supervised learning allows the researcher to define the categories in which one is interested. Supervised learning is therefore best when one knows for what one is looking, such as tweets coordinating protest or talking about a state's repressive response. See Grimmer (2013) and Lucas (2015) for more detail on text analysis.

After deciding upon a supervised topic model, a grid search was performed to tune model parameters. Specifically, a support vector machine and 3 varieties of Naive Bayes classifiers were tested. Each were tried with varying numbers of document features, with the features defined 1, 2, or 3 n-grams. Those combinations were then trained on 95% of the coded tweets and tested on the remaining 5%; the specific 95% was chosen at random with replacement. The training was repeated a variable number of times, from once to fifty, and a tweet is classified as belonging to a category if more than half of the models assign it to that category. This process is known in machine learning as a bagged ensemble, is equivalent to bootstrapping, and lowers the variance of a model. This entire process was performed 4 times: once on protest coordination tweets in Egypt and Bahrain, and then again for common knowledge ones in each country. The model with the highest F1 score was chosen and used to generate predictions for the out of sample data. The final classification parameters are shown in Table 2.2.

Table 2.2: Topic Models' Parameters

Country	Topic	Classifier	Features	N-grams	Bags	F1
Bahrain	Coordination	Bernoulli Naïve Bayes	450	3	30	.65
Bahrain	Common Knowledge	Bernoulli Naïve Bayes	600	1	15	.62
Egypt	Coordination	Bernoulli Naïve Bayes	800	2	25	.64
Egypt	Common Knowledge	Bernoulli Naïve Bayes	900	3	10	.59

Other measures are included that may capture coordination. For example, a retweet repeats information, meaning that a day with many retweets may lead to more common knowledge. Tweets with user mentions reflect direct engagement between accounts that is visible to more than just the tweet author and mentioned user; it shows that the information expressed is seen by at least one other person. Increases in the proportion of tweets with links or hashtags should have similar effects. Knowing that many people are talking about policy disagreements, that particular tweets, some of which are about grievances or upcoming protests, are retweeted and so widely seen, or that many tweets contain user mentions, suggesting a grievance is shared by more than just the tweeter, all may contribute to coordination. Measures of the percent of tweets which contain retweets; the percent which contain user mentions; the percent which have at least one hashtag; and the percent which contain a link therefore capture coordination which may occur outside of the Gini coefficient.

Finally, I control for state repression and previous levels of protests. While the conditional effect of repression on protest is an active area of research [Moore, 1998, Carey, 2010, Siegel, 2011], the two are clearly related [Davenport, 2007*a*]. Since protests are a dynamic process where the level of protest on one day can affect the level the next [Lohmann, 1994, Rasler, 1996], controlling for a previous day's number of protests is necessary. Both measures are taken from ICEWS.

Because the dependent variable is a non-negative count with an unequal mean and variance, a negative binomial model is used to model the relationship between the independent variables and protest. The base model is:

$$Protests_{i,t} = \beta_0 + \boldsymbol{\beta_1} * \boldsymbol{\Omega_{i,t-1}} + \boldsymbol{\beta_2} * \mathbf{X}_{i,t-1} + Protests_{i,t-1} + Repression_{i,t-1} + \epsilon_{i,t} \quad (2.1)$$

where Ω represents the independent variables of interest in each model, $\mathbf{X}$ represents a series of controls, and ϵ is a stochastic error term. Since $Protests_{i,t}$ is overdispersed and the zeroes should be true zeroes, a negative binomial model instead of a Poisson or zero-inflated negative binomial is used. All independent variables are lagged one day, country fixed-effects are used, and standard errors are clustered by country.

2.5 Results

H1: Common Knowledge

The start of mass protests in Egypt on January 25^{th}, 2011 was the culmination of the common knowledge creation process started by activists. Mohamed ElBaradei, the Egyptian former head of the International Atomic Energy Agency who won the Nobel Peace Prize in 2005, had returned to Egypt in February 2010 and immediately became a focal point of anti-Mubarak organization; he formed a group, the National Association for Change, which gathered millions of signatures demanding a list of far-reaching reforms, the most important of which was the cessation of the Emergency Law [Khalil, 2011, pgs.106-119]. Many of the core members who collected signatures and wanted to use the NAC for challenging Mubarak were disappointed when ElBaradei did not press his momentum, and many of these people would be the same ones organizing events a year later [Khalil, 2011, Cambanis, 2015, pg. 111; pg. 46]. Common knowledge that large groups of individuals had mutual policy disagreements therefore was increasing throughout 2010, to say nothing of the decade's previous labor and internationally-oriented activism [Gunning & Baron, 2013].

Two events in advance of January 25^{th} further contributed to common knowledge. On June 6^{th}, 2010, Khaled Said was attacked by two police as he entered an internet cafe in Alexandria. As he was beaten, the police dragged him back into the street and into the foyer of a nearby building, where passersby watched the officers beat him to death [Khalil, 2011, pgs. 69-87].[4] At the morgue the next morning, a family member took a photo of his wrecked face, and this photo quickly spread across Egypt via cell phone messages and via a Facebook group, "Kullena Khaled Said" (We Are All Khaled Said). This page was notable for two reasons. First, it was explicitly designed to have a broad appeal; it was not the first to draw attention to Said's death, but it was the first to not explicitly associate itself with anti-regime activists. The

[4]The two officers were not detained until June 30^{th}, after the photo spread nationwide and the government stridently denied responsibility. After a series of delays, the trial started on October 22^{nd}, 2011, and the two officers were quickly found guilty of manslaughter and sentenced to seven years in prison. That sentence was increased to ten years in 2014.

administrator, Wael Ghonim, remained anonymous, and the page quickly became a forum for organizing offline events [Ghonim, 2012, pgs. 58-81].[5] Second, the page quickly gained hundreds of thousands of followers, and these followers openly shared their grievances and discussed action to take.[6] The page made it clear that there was a large reservoir of anti-regime sentiment.

Wael Ghonim established "We Are All Khaled Said" on Facebook on June 10^{th}, 2010. From the beginning, the page was a fulcrum of free expression, with most of that expressing rage against Mubarak and his regime. It was used to organize a series of "Silent Stands" across Egypt, events where individuals would stand, dressed in all black, near each other, forming long chains of people. Having attracted thousands of people, the events were an early demonstration of increasing dissatisfaction and foreshadowed the role the Facebook page would play in future events.

Ghonim also strategized and coordinated with other activists, including those of the April 6^{th} movement [Ghonim, 2012, pg. 114, 121]. In late December, these activists decided to organize an event for January 25^{th}, 2011, National Police Day. The police were perhaps the most hated institution in Egypt; the holiday was started in 2009 to honor them, and it was already used as a protest day by the April 6^{th} group in 2010. Once President Ben Ali fled Tunisia on January 14^{th}, Ghonim and his close associates used the page to reach out to other groups and encourage them to protest. The idea soon gained traction, and other groups started advertising the upcoming protest [Ghonim, 2012, pgs. 139-141].

Ghonim and his close associates understood that a Facebook page could not catalyze a nation. Internet penetration was (and is) low in Egypt, and Facebook

[5] Ghonim made his identity known to a select group of confidants, including his co-administrator, AbdelRahman Mansour.

[6] 365,000 followers by early 2011, according to one account [Cambanis, 2015, pg. 46].

was an even smaller percentage of that. The page's members were clearly more educated and wealthier than most Egyptians, and a message that stayed inside those social circles would not lead to the revolution many sought.[7] To avoid this fate, administrators asked the page's members to use text messages, fliers, and in-person communication to raise awareness of the upcoming protest [Ghonim, 2012, pg. 143].

The second crystallizing event was the bombing of a Coptic church in Alexandria in the first minutes of 2011. The bombing, which killed 21, first angered Christians who felt the state, their protector from Muslim violence, broke their social contract; the government's haphazard roundup and torture of Salafists then alienated many non-Christians as well [Khalil, 2011, pgs. 183-186]. The subsequent dissatisfaction was now spread across disparate groups with little else in common, and it was after this bombing that activists began to plan intensely for upcoming action. At this point, the "We Are All Khaled Said" page was active, but activists continued to rely on in person, clandestine meetings to discuss tactics and strategy [Cambanis, 2015, pgs. 46 - 47]. The "We Are All Khaled Said" was once again used to generate common knowledge that policy dissatisfaction was widespread in society.

As January 25^{th} approached, preparatory activity increased. Activists continued to meet offline, and Ghonim used his (anonymous) online persona to guide the Facebook page; at this point, thousands of new members joined every day [Ghonim, 2012, pg. 153]. Ghonim posted several long messages enumerating the grievances around which everyone could rally. He would eventually use the group to advertise a comprehensive Google Doc that served as a manifesto: it explained who the group was, why protest was justified, demands, information on when and where to protest,

[7] Later survey research of protesters at Tahrir confirmed the non-representativeness of social media users in Egypt [Tufekci & Wilson, 2012].

behavior guidelines, and chants to use during marches [Ghonim, 2012, pgs. 164 - 169].

Though the actual day of January 25^{th} is seen as the start of the major protests in Egypt, it is better to think of it as the end of the common knowledge generation process. As explained in the theory section, protests do not become revolutionary until "normal" people observe initial protests and see people that look not much different from them out protesting [Lohmann, 1994]. For many reasons, it was not clear on January 24^{th} exactly what kind of protest the next day would entail. State media was claiming that protesters were foreigners, and many organizers were happy that the Muslim Brotherhood did not mobilize its base initially, as that would have made it easy for the government to paint the protesters as religious radicals [Khalil, 2011, pgs. 137-138]. Protests started at many different locations on Cairo's periphery, and organizers exhorted bystanders to join them as they moved towards Tahrir Square, the center of Cairo [Cambanis, 2015, pg. 51]. During the morning and early afternoon, many early participants were still very active on social media, including Twitter, to convey protest locations, the number of protesters, and urge others to join [Khalil, 2011, Idle & Nunns, 2011, pg. 139; 33-36].

H2: Coordination

Table 2.3 provides regression results to support the argument that activists do not affect protests through coordination. The control variables for coordination - the percent of tweets from activists, the percent of tweets that are activist retweets, the percent of tweets that are activists mentioning others, and the percent of tweets with hashtags that are from activists, and the percent of tweets with links from activists

- do not correlate with subsequent protest. *Activist Coordination Tweet* $\%_{i,t-1}$, the percent of tweets about coordination that are from activists, negatively correlates with protest. The intuition here is that activists are always trying to coordinate protests but non-activists only sometimes do; since more protests occur when non-activists coordinate protests [Steinert-Threlkeld, 2016], large values of *Activist Coordination Tweet* $\%_{i,t-1}$ correspond to periods of low protest.

The 2nd model in Table 2.3 uses a different measure of activist coordination but is otherwise identical to the 1st. Here, *Activist Coordination*$_{i,t-1}$ is the interaction of Non-Activist Coordination$_{i,t-1}$ and Activist Hashtag $\%_{i,t-1}$. Non-Activist Coordination$_{i,t-1}$ is the Gini coefficient for hashtags per country-day; if a large percentage of those hashtags are produced by activists, then hashtag coordination comes from activists. This measure is not statistically significant, while *Activist Coordination Tweet* $\%_{i,t-1}$ remains so. Note that in both models Non-Activist Coordination$_{i,t-1}$ is significant with a p-value less than .01. Because model 2 does not change results, model 1 from Table 2.3 is the main model used for subsequent robustness checks.

Table 2.3: Main Results

	DV: Protest$_{i,t}$	
	(1)	(2)
Activist Tweet %$_{i,t-1}$	2.582	1.718
	(6.978)	(7.472)
Activist Retweet %$_{i,t-1}$	5.649	6.256
	(5.349)	(5.541)
Activist Mention %$_{i,t-1}$	−2.233	−1.155
	(7.320)	(8.077)
Activist Hashtag %$_{i,t-1}$	−1.220	−1.275
	(1.133)	(1.137)
Activist Link %$_{i,t-1}$	9.751	11.187
	(11.805)	(12.182)
Protest$_{i,t-1}$	0.014	0.014
	(0.008)	(0.008)
Repression$_{i,t-1}$	0.007	0.007
	(0.018)	(0.018)
Non-Activist Coordination$_{i,t-1}$	5.314***	5.329***
	(0.847)	(0.848)
Activist Coordination Tweet %$_{i,t-1}$	−253.024**	−218.724*
	(116.865)	(132.533)
Activist Coordination$_{i,t-1}$		−181.106
		(393.024)
Intercept	−2.062***	−2.081***
	(0.494)	(0.496)
Country FE	Yes	Yes
N	168	168
Log Likelihood	−473.423	−473.333

*p < .1; **p < .05; ***p < .01
NB: All activist variables except for Activist Hashtag %$_{i,t-1}$ use the total tweets from country$_i$ as the denominator. Activist Hashtag%$_{i,t-1}$ uses the total tweets with hashtags from country$_i$. Section 2.4 explains this modeling decision.

Note that *Activist Hashtag* %$_{i,t-1}$ is calculated slightly differently than *Activist Tweet* %$_{i,t-1}$, *Activist Retweet* %$_{i,t-1}$, *Activist Mention* %$_{i,t-1}$, and *Activist Link* %$_{i,t-1}$. *Activist Hashtag* %$_{i,t-1}$ is calculated as the percent of all tweets with hashtags that are tweets from activists, but the other 4 take the total number of tweets from the activists' country on that day as the denominator. The variables are modeled differently to reflect the information consumption process on Twitter. When one sees a tweet on Twitter, it is presented as part of a sequence of reverse chronological tweets. If one views tweets containing a hashtag, however, all tweets in the subsequent reverse chronological sequence contain that hashtag. The determinant of the length of the latter sequence is therefore all tweets containing that hashtag while the length of all tweets one sees is better approximated by all tweets on that day.

Two sets of empirical checks validate the results in Table 2.3. Table 2.4 replicates model 1 from Table 2.3 but with country robust standard errors and using count variables. In the first model, country robust standard errors are run. Robust standard errors were not used originally because the data only have two clusters, Egypt and Bahrain, and clustered standard errors with few groups can bias standard errors down [Donner, 1998, Wooldridge, 2003, Arceneaux & Nickerson, 2009]. Using country-clustered standard errors increases the standard error of *Activist Coordination Tweet* $_{i,t-1}$, making it no longer statistically significant. *Activist Hashtags*$_{i,t-1}$ now negatively correlates with subsequent protest, and the lagged dependent variable is statistically significantly correlated with protest.

Though using counts instead of proportions is a less accurate model of how information is consumed on Twitter, the models in Table 2.3 were rerun using count variables. These results are Model 2 of Table 2.4. Using count variables, activist coordination negatively correlates with subsequent protest. The count of hashtags does positively correlate with protests, which is not surprising since activists, like everyone else, tweets more during protest events. In both models, the level of significance and signs do not change for *Protest*$_{i,t-1}$ and *Non-Activist Coordination*$_{i,t-1}$.

Table 2.4: Robustness Checks - Clustered SEs, Count IVs

	DV: $\text{Protest}_{i,t}$	
	IV: Percentage	IV: Count
	(1)	(2)
Activist $\text{Tweets}_{i,t-1}$	2.582	-.0.005
	(9.875)	(0.004)
Activist $\text{Retweets}_{i,t-1}$	5.649	-0.004
	(5.751)	(0.003)
Activist $\text{Mentions}_{i,t-1}$	−2.233	0.002
	(10.645)	(0.004)
Activist $\text{Hashtags}_{i,t-1}$	−1.220*	0.007*
	(0.719)	(0.004)
Activist $\text{Link}_{i,t-1}$	9.751	0.004
	(8.986)	(0.005)
$\text{Protest}_{i,t-1}$	.014**	0.017*
	(0.007)	(0.010)
$\text{Repression}_{i,t-1}$	.007	0.002
	(0.024)	(0.018)
Non-Activist $\text{Coordination}_{i,t-1}$	5.314***	4.927***
	(1.733)	(0.780)
Activist Coordination $\text{Tweets}_{i,t-1}$	-253.024	-0.067***
	(290.352)	(0.023)
Intercept	−2.062***	-1.703***
	(1.093)	(0.424)
Country FE	Yes	Yes
Clustered SE	Yes	No
N	168	168
Log Likelihood	−472.423	−471.326

*p < .1; **p < .05; ***p < .01
NB: Model 1's variables are calculated as in model 1 of Table 2.3. Model 2's are counts.

Finally, I rerun the models including a measure for activist tweets from mobile devices. Activists in both countries frequently use mobile devices for communication, but they also use desktop computers. I measure the switching between devices by taking advantage of a key piece of metadata that comes with tweets, the "tweet_source" field. This field is a string created by Twitter to reflect the provenance of a tweet.[8] The results do not change. The percent of an activist's tweets that are from a mobile

[8]The Supplementary Materials provides more detail on modeling mobility and how mobility patterns changed during the course of protests.

phone does positively correlate with subsequent protest and has a p-value between .05 and .10, though the result does not hold when that value is interacted with the percent of tweets that are about coordination. This result makes sense, as a tweet does not say whether or not it comes from a mobile device, so a tweet from one does not provide a signal to others that the author has mobilized.

Table 2.5: Robustness Checks - Tweets from Phones

	Protest	
	(1)	(2)
Activist Tweet $\%_{i,t-1}$	-6.285	-6.138
	(8.062)	(8.053)
Activist Retweet $\%_{i,t-1}$	-1.362	-1.911
	(6.848)	(6.903)
Activist Mention $\%_{i,t-1}$	8.400	8.024
	(8.989)	(9.004)
Activist Hashtag $\%_{i,t-1}$	-0.845	-0.502
	(1.111)	(1.099)
Activist Link $\%_{i,t-1}$	10.904	7.935
	(11.708)	(11.735)
$\text{Protest}_{i,t-1}$	0.015^{*}	0.013
	(0.008)	(0.008)
$\text{Repression}_{i,t-1}$	0.009	0.014
	(0.018)	(0.018)
Non-Activist $\text{Coordination}_{i,t-1}$	5.444^{***}	5.598^{***}
	(0.844)	(0.846)
Activist Coordination Tweet $\%_{i,t-1}$	-271.442^{**}	-437.983^{***}
	(116.122)	(168.530)
Mobile Phone $\%_{i,t-1}$	6.043^{*}	5.440
	(3.516)	(3.520)
Activist Coordination Tweet $\%_{i,t-1}$* Mobile Phone $\%_{i,t-1}$		4,767.435
		(4,194.302)
Intercept	-2.094^{***}	-2.166^{***}
	(0.491)	(0.493)
Country FE	Yes	Yes
N	168	168
Log Likelihood	-472.111	-471.262

*p < .1; **p < .05; ***p < .01

The Supplementary Materials explores the tweets in more detail. It shows that the hashtags which were most important to activists were hardly used outside of the activist network, while hashtags that non-activists used quite frequently were barely used by activists. Activists may not influence protests through coordination because they do not talk about topics of interest to non-activists. This disconnect may affect the ability of activists to coordinate.

2.6 Activists and Policy Negotiation

Activists are busy during protest events but do not have a coordinating effect. What, then, is their role?

They engage in behaviors which benefit when only a few people engage in them, such as negotiation with the current regime. Determining a policy platform and negotiating over that platform with authorities is activity best handled by a small group, as involving everyone in policy drafting and negotiation is likely to lead to no platform being drafted, a drafted platform too broad to be actionable, or consultation too slow to facilitate negotiation with a regime. Disarray and torpor are especially harmful, as a government can wait out protesters and some protesters will become violent if negotiations appear to stall. Indeed, activists in Morocco failed to agree on policies to propose to King Mohammed VI, limiting their ability to lead the ongoing protests [Benchemsi, 2014]; after Egypt's 18 days of protest, the core activist leaders would face the same problem and see their efficacy diminish [Cambanis, 2015, pgs. 101-108]

In other words, activists become leaders of a pseudo-government that presides over a protest event. This paper makes no prediction about how they become leaders - there could be elections within a protest site, or the activists could be chosen based on past experience, for example. Once protests are underway, there is no need for coordination; that activity belongs to non-activists. But non-activists are too numerous and inexperienced to negotiate with authorities for a policy change, and authorities will find it more difficult to negotiate with a mass of individuals than with a small group. The activists lead, but protest mobilization is not within their purview.

First-hand accounts of protests in Egypt corroborate this argument. Activists quickly assumed leadership roles once protests started; the leadership first focused on street-level negotiations with police commanders and then shifted to higher-level negotiations once protesters were firmly in control. In clashes with police, both on the 25^{th} and 28^{th}, street-level negotiations would occur between protest leaders and police commanders. Though there is no evidence that these negotiations led to less violence or a change in tactics, they reveal that those who were key common knowledge creators were doing more than just marching [Khalil, 2011, pgs. 148,174]. As soon as the square was occupied, various activists found each other and began to work together on political platforms, protests logistics, and recruitment [Cambanis, 2015, pgs. 60-62]. On February 2^{nd}, 14 activists ranging from the Muslim Brotherhood to the socialist left, formalized their activity by creating the Revolutionary Youth Coalition (RYC). Each member had veto power and was responsible for conveying the main interests of the group they represented. When the regime tried to negotiate with protesters, it was by negotiating with this body, but the RYC would not meet with government officials until Mubarak resigned [Cambanis, 2015, pgs. 60-69]. In fact, the RYC formed the day after a meeting with Vice-President Omar Suleiman, also the head of the General Intelligence Directorate [Ghonim, 2012, pgs. 240-241].[9]

Other activists did meet with regime officials. In the bus ride from prison to his family, on February 7^{th}, Wael Ghonim was joined by Mahmoud Wagdy, briefly the Interior Minister, and Dr. Hosam Badrawy, a senior member of the National Democratic Party [Ghonim, 2012, pgs. 251-252]. The next day, the Minister of Information, Anas El-Feky, asked Ghonim, through his brother, to appear on state

[9] The order to release Wael Ghonim from prison may have resulted from this meeting.

television; having said no, El-Feky personally appealed to Ghonim. The night of February 9th, after an interview on CNN, the regime, through Badrawy and another intermediary, again reached out to Ghonim, and Ghonim met with the Minister of the Interior until 3 a.m. Less than 6 hours later, Ghonim met with Dr. Badrawy again; Dr. Badrawy claimed to have a direct line to Mubarak and promised a direct meeting with the president to discuss his resignation [Ghonim, 2012, pgs. 270-276]. A long meeting with lower level ministers made it clear that meeting would never happen, and that meeting appears to be the final negotiation between activists and the regime [Ghonim, 2012, pgs. 275-282]. Nowhere in his accounts of this period does Ghonim mention Twitter, and he appears to have worked through the "We Are All Khaleed Said" page much less than before his arrest.

The night of February 10th, Mubarak would make another speech without resigning; on the 11th, Vice-President Suleiman announced Mubarak had handed power to a committee representing the armed forces. At this point, many protesters believed their main goal was accomplished, and activists found themselves with fewer people to lead [Cambanis, 2015, pg. 84]. Many of the activists believed, however, that Mubarak's abdication was only the start of a policy negotiation process and did not trust the Supreme Council of the Armed Forces (SCAF), the new, self-declared governing body, to operate in the best interests of Egyptians. These activists therefore had to switch back to a common knowledge creation role; many would soon appear on national talk shows and travel the country trying to convince people to protest again [Cambanis, 2015, pgs. 85-90]. A detailed description of the rest of 2011 is outside the scope of this paper, but events, and especially the SCAF and Muslim Brotherhood, would soon marginalize the activists in this study.

2.7 Conclusion

Activists translate dissatisfaction with policy into protests, with the goal of changing the status quo policy. Because many countries provide few channels for most individuals to express policy preferences, protests become a primary vehicle for policy negotiation. Activists try to convince non-activists that their anti-regime preferences are widespread. The activists then must convince individuals to protest. With protests underway, activists undertake two activities: coordinating the protests and negotiating for policy change. Coordination entails directing protesters where to go, when to go there, how to behave, and what supplies to bring. Negotiation includes trying to convince police commanders to not resist protesters, drafting policy platforms, and presenting those platforms to those holding onto power. Activists are effective through common knowledge generation and policy negotiation but not through coordination.

This paper has shown that social media data can be used to understand events occurring outside of social media, but it does not argue that social media was essential to the Arab Spring. Social media have become a tool activists and protesters have used [Tufekci & Freelon, 2013, Valenzuela, 2013], but it is not a silver bullet and probably did not cause the Arab Spring (or Iran's 2009 post-election protests, or Turkey's Gezi Park protests, or Brazil's Vinegar protests, or the #Euromaidan events, and so on). Using social media data to understand social media's impact on offline events requires precise geolocation data to provide variation on social media prevalence across locations, and those data are few except in advanced industrial countries and a few large middle-income countries (Brazil, Turkey, and Indonesia).

While the evidence presented via Twitter and memoirs appears to support the claim that activists affect protest through common knowledge and policy negotiation, 3 alternative explanations that may explain the patterns observed. First, the activists studied here might not have been the right group to study for protest mobilization. In fact, while there were certainly other groups and individuals working against the regime, most bandwagoned. Not until February 1^{st} did the judges' association publicly support the protest; Mohammed ElBaradei did not organize protests and did not visit Tahrir Square until January 30^{th}; and Omar Sharif, a popular actor, publicly sided with protesters on the 1^{st} [Khalil, 2011, pgs. 212-213].[10] Individual journalists started to chafe at the talking points handed to them, but they were either fired or had no effect [Khalil, 2011, pgs. 284 - 287]. Not until Wael Ghonim's interview with Mona El-Shazly on the evening of February 7^{th} did state-owned media show its viewers an unvarnished representation of the protests [Ghonim, 2012, pgs. 254-258]. Until then, the protesters were supposedly English-speaking foreigners paid for by the Israeli and American governments.

Al-Jazeera was a key source of independent information and reported intensely on the protests as they unfolded. But the channel was not involved in creating common knowledge before the start of protests, and it quickly suffered heavy repression once the engagement phase started. Its journalists were put in jail or harassed, and its license was quickly revoked [Khalil, 2011, pgs. 251-255]. It was an important conduit of information, especially in the initial days around January 25^{th}, but it was not a key player before protests began.

The activists also may be the wrong unit to study because they did not have the scale to challenge the state and replace the regime. The only organization with a

[10] Indeed, many celebrities sided with Mubarak throughout the protests.

nationwide network, experienced leadership, and a political platform was the Muslim Brotherhood. While the Brotherhood provided significant logistical support and manpower after January 25^{th} - their announcement of their intention to join the protests on January 28^{th} was a major turning point - they were not involved in preparing for the first protests. Indeed, early protesters were quick to say they were not part of the Muslim Brotherhood, and the non-religious character of the January 25^{th} protests was evident. Moreover, disagreement existed between Brotherhood leadership and many of its youth, the latter preferring the direct preparation and engagement associated with the activists in this study [Cambanis, 2015]. The Brotherhood was instrumental for the revolution once it got underway, but it did not create common knowledge. The extent to which it negotiated with the regime once protests began is not known.

Second, it is possible that social media data cannot provide insight into outcomes outside of social media. They do. From mood patterns [Golder & Macy, 2011] to the flu [Chew & Eysenbach, 2010], natural disasters [Starbird & Palen, 2010] to political events [Starbird & Palen, 2012], behaviors on Twitter clearly reflect offline activity. Twitter can even be used to detect upcoming events [Garcia-Herranz et al., 2014], including changes in the stock market [Bollen, Mao & Zeng, 2011] or sports outcomes [Sinha et al., 2013]. While Twitter users in Egypt are not representative of non-Twitter users [Tufekci & Wilson, 2012], it is not clear that that difference bias protest-related behavior observed on Twitter. Communication and differential influence also occur offline; social media do not change the structure of social relations, the importance of widespread grievances and changing political opportunity structures, or the need for offline leadership [Tufekci, 2014]. The main benefit of social media is therefore to lower the cost of data gathering. The lower cost of data gathering benefits anyone who uses social media: individuals who coordinate protests, state authorities who monitor those events [Bassiouni et al., 2011, Gunning & Baron, 2013], and academics who study either. For academics, social media therefore provide a window into protest dynamics, shedding some light into what had previously been a black box [Segerberg & Bennett, 2011, Bennett & Segerberg, 2013].

Third, it is possible that coordination was not observable using Twitter because of fear of state repression. That states use social media to target individuals for repression is sometimes observed and less often studied [Bassiouni et al., 2011, Gunitsky, 2015], and it could cause the decrease in Twitter activity observed here. Fear of repression does not appear to drive these results, for three reasons. First, if activists were afraid of using their phones, there would be a decrease in the percentage of tweets coming from cell phones, which did not happen (see the Supplementary Materials). Second, nowhere in the memoirs cited throughout this paper do activists explain their decrease in tweeting is a result of fear of repression; of anything, they were more fearful before the protests' start, as it was quite common to remove SIM cards and batteries before important meetings [Cambanis, 2015, pg. 46]. It was also common for activists to tweet their locations during protests as a safety measure, as it would then be clear if the state captured them and where [Ghonim, 2012]. Third, Egypt's nationwide blackout suggests the government lacked the power or desire to narrowly repress based on cell phone behavior. Narrow repression - silencing the well-connected voices - is a superior repression strategy than its indiscriminate cousin [Rasler, 1996, Siegel, 2009], so the mass suppression of cellular communication may have suggested to activists that they were not going to be directly targeted.

This paper has demonstrated the benefits of combining online and offline data.

Little work focusing on social media connects patterns there to offline activity, while few case studies and ethnographies analyze online behaviors in a quantified manner. Combining quantified social media data with qualitative data on contemporaneous offline activity therefore provides a more complete picture of an activity than keeping the two data types separate. Social media analysis provides the forest and case studies taxonomize the trees, but rarely do the two combine.

Future work should analyze activists' behavior when protests do not enlarge. During the Arab Spring, Saudi Arabia, Jordan, Algeria, Iraq, and Morocco featured protest movements that did not become self-sustaining the way they did in Tunisia, Syria, Libya, Egypt, or Bahrain. In both sets of countries, states responded with repression, economic stimulus, or a combination of the two. A key difference in their outcomes may therefore be found in the tactics undertaken by those countries' activists.

2.8 Supplementary Material

2.8.1 Detail on Topic Models

Each tweet was coded for one of 39 categories, and a tweet could belong to multiple categories. For the topic model, these categories are aggregated to 8: protest coordination, common knowledge, protest support, the state's response, anti-protest, foreign protests, religious tweets, and tweets about events happening the next day. Protest support is for any tweet that is about protest coordination, protest information, corruption, democracy, events happening tomorrow, or contains anti-regime sentiment. The anti-protest category is for a tweet that is anti-protest or pro-regime. Every tweet that is about Arab Spring events is grouped into the foreign protests category.

Python's Natural Language Toolkit (NLTK) prepared the 3,000 tweets for analysis. For text analysis, it is important to normalize each document, which means performing certain operations on each word to remove idiosyncrasies across documents. Normalization includes removing stopwords (common words such as "and", "the", "is", that are so common they do not distinguish documents), lemmanizing words (converting related words to their common stem, e.g. "stopped", "stopping", and "stoppage" all become "stop"), converting all words to lowercase, and removing punctuation and symbols.[11] NLTK provides stopwords for English; for Arabic, I borrowed a custom stopword dictionary of 275 words used to code Facebook posts of the April 6 movement [Hanna, 2013]. Normalizing each training set means that the content of each tweet is removed of enough idiosyncrasies so as to provide enough information to the support vector machine to allow it to make predictions. The dataset where each tweet is assigned to a predicted category (or categories) is the final dataset. Sample tweets for each category are shown in Table 2.6.[12]

After cleaning the tweets, Python's scikit-learn library is used to build the support vector machine. In text classification, the variables are words, and the number of variables is equal to the number of unique words found across all documents (tweets). Two problems arise: common words will be overrepresented, and there will be thousands of variables, increasing computation time and the possibility of overfitting. To counteract these, I build a Term Frequency Inverse Document Frequency matrix, which creates weights for each word based on how often it appears and in

[11] I did not remove the @ symbol because it is used in retweets and for coordination.

[12] @7areghum is well-known for being a pro-government account in Bahrain and may even be run by the government [Bassiouni et al., 2011].

Table 2.6: Tweets by Category

Protest coordination		
March start @ 3PM to roundabout	02.25.2011	@byshr
Common Knowledge		
RT @RamyRaoof: Aggregating Chants by People of #Egypt in Peaceful Assemblies http://egychants.tumblr.com [Ar] #Jan25	01.23.2011	@monasosh
Protest support		
We are more powerful than the dictator. "The ruler can only rule with the consent and cooperation of the ppl" Robert Helvey	02.13.2011	@angryarabiya
State Response		
#Bahrain tear gas and rubber bullets used against peacful protester this morning in nuwaidrat #Feb14 #Arab #Humanright #Torture	02.14.2011	@saidyousif
Anti-Protest		
Oh merciful god that Bahrain and its Sunni people love, oh god, if this was a test of understanding, we will persevere	03.03.2011	@7areghum
Foreign protests		
@anwarshaikh46 this is not egypt, but a dictator is a dictator. Egypt is not Tunisia either, but being inspired by them they had a victory	02.22.2011	@angryarabiya
Religious		
GroupWednesday night prayer beseeching God bless your parents	03.22.2011	@14febrevolution
Next day		
@arabist: new tunisia govt tomorrow (i hear min finance interior gone) rt @arouabensalah: #ttn l'annonce du nouveau gouvernement demain	01.25.2011	@alaa

how many documents; low weights are given to frequent words in many documents, high weights to frequent words in few documents.

Table 2.7 is the same as Table 3 from the main paper except the precision, recall, and AUC scores are reported in place of the F1 score. Precision means the percent of all tweets labeled by the classifier as belonging to Topic X that actually belong to Topic X; it is the true positive rate. Recall is the percent of all tweets about Topic X that are identified by the classifier. The AUC curve plots the trade-off between the true positive and true negative rate of a model, and the AUC score is the area under that curve; a score of 1 means a high true positive rate is obtained without identifying any false negatives.

Table 2.7: Topic Models' Parameters

Country	Topic	Classifier	Features	N-grams	Bags	Precision	Recall	AUC Score
Bahrain	Coordination	Bernoulli NB	450	3	30	.83	.53	.76
Bahrain	Common Knowledge	Bernoulli NB	600	1	15	.68	.57	.74
Egypt	Coordination	Bernoulli NB	800	2	25	.62	.67	.82
Egypt	Common Knowledge	Bernoulli NB	900	3	10	.62	.57	.77

2.8.2 Mobility

I also measure the percent of tweets which come from a mobile device. Activists in both countries frequently use mobile devices for communication, but they also use desktop computers. I measure the switching between devices by taking advantage of a key piece of metadata that comes with tweets, the "tweet_source" field. This field is a string created by Twitter to reflect the provenance of a tweet. In this dataset, a tweet is classified as mobile or not according to the following mapping:

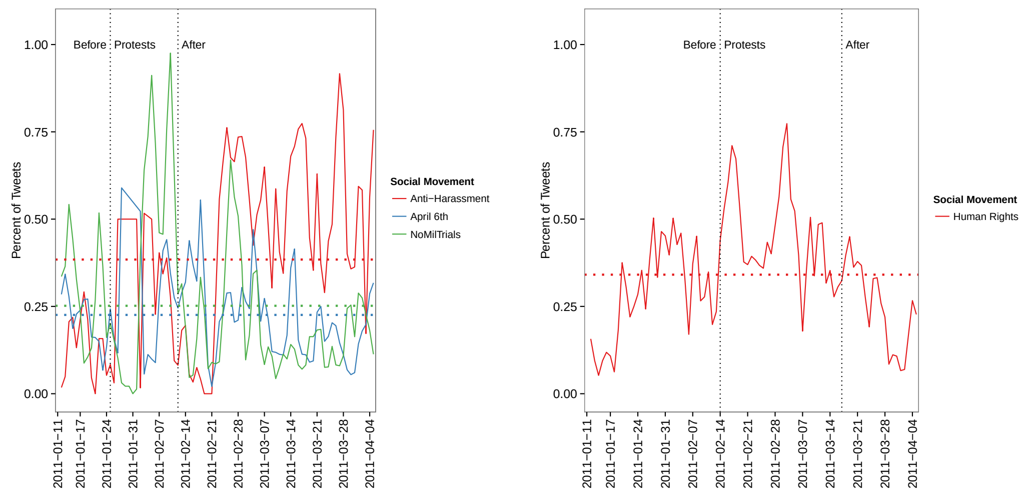
- **Mobile:** 'Twitter for BlackBerry©'; 'Gravity'; 'Tweet Button'; 'twitterfeed'; 'Twitter for iPhone'; 'Facebook'; 'My Tweet Lovers'; 'harassmap.org'; 'StumbleUpon iPhone'; 'Mobile Web'; 'oauth:173069'; 'oauth:3294'; 'Twitpic'; 'Bambuser'; 'Google'; 'Twitter for iPad'; 'Snaptu'; 'ÜberSocial'; 'Samsung Mobile'; 'TweetMeme'; 'Yfrog'; 'TwitLonger Beta'; 'See Wh0 Viwed Y0ur Pr0file'; 'The BOBs'
- **Not Mobile:**[13] 'Web'; 'Choqok'; 'TweetDeck'; 'HootSuite'; 'Ping.fm'

These metadata are not provided in the data from Mocanu et. al, so mobility cannot be computed for non-activists.

Figure 2.1 presents the changing percentage of activist tweets from mobile devices in Egypt and Bahrain. The y-axis shows the two day rolling average of the percent of each movement's tweets which come from a mobile device, the x-axis shows

[13]HootSuite and TweetDeck have mobile versions. Since they were originally designed for desktop use, I assume that tweets from them are not from mobile phones.

the date. In both countries, more tweets come from mobile devices once protests start, though the relationship with protests is not always statistically significant.

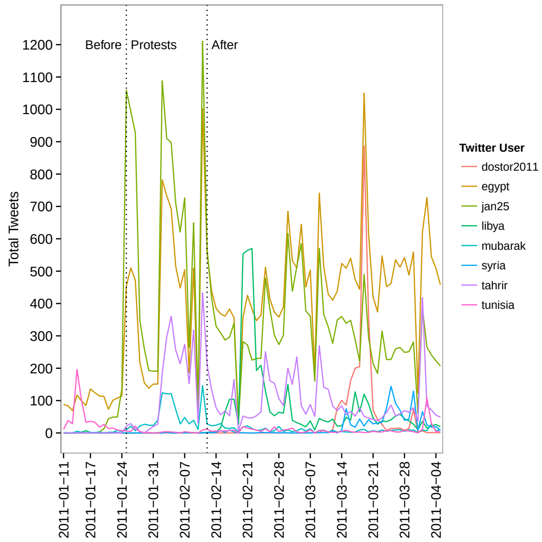


(a) Egypt (b) Bahrain

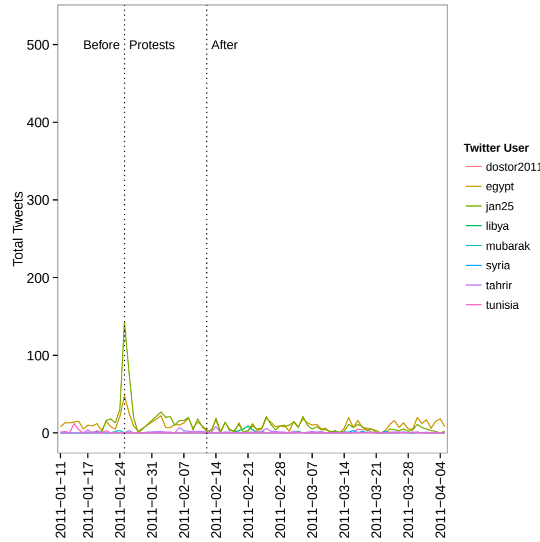
Figure 2.1: Mobility in Egypt, Bahrain

2.8.3 Hashtag Topics by Social Movement

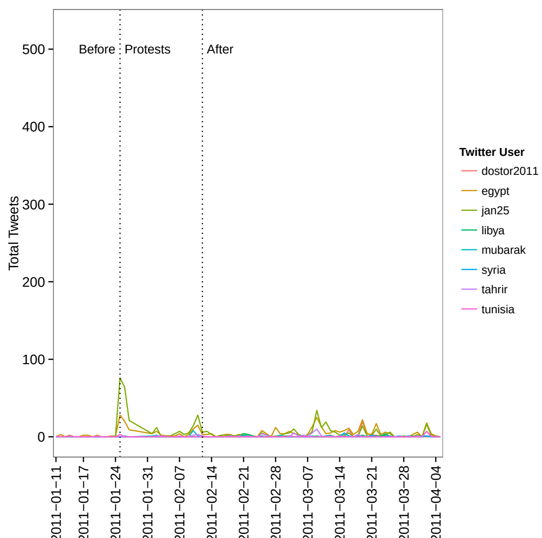
This section presents the common and unique hashtags per movement in Egypt and Bahrain.



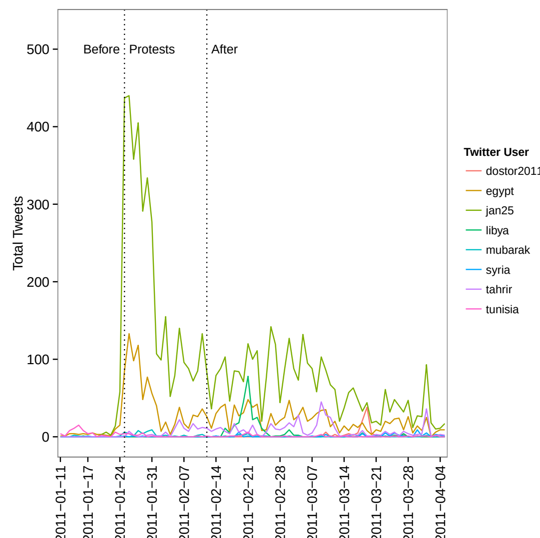
(a) Non-Activist Egyptians



(b) April 6^{th}

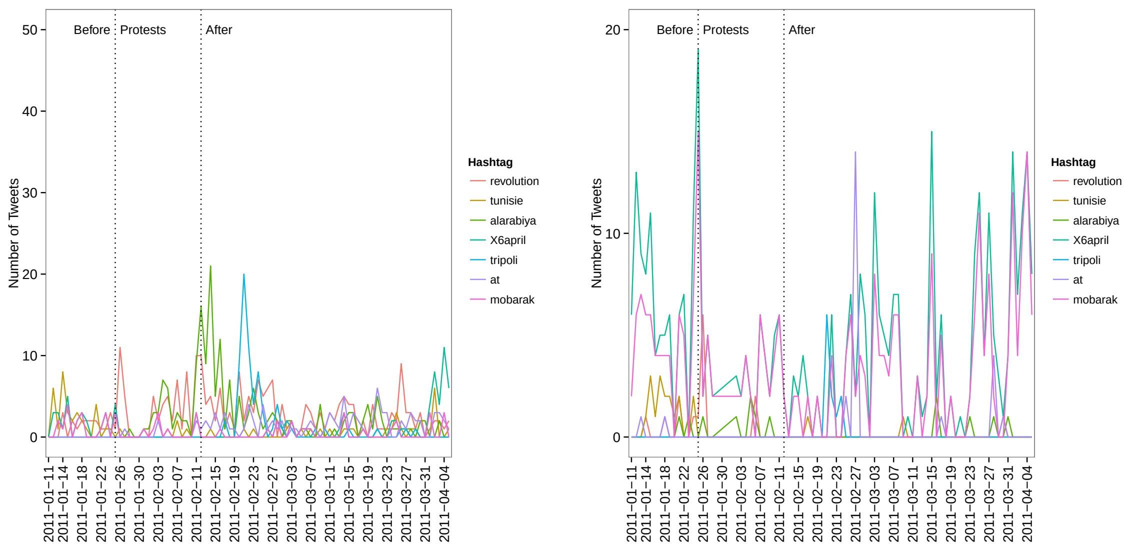


(c) Anti-Sexual Harassment



(d) NoMilTrials

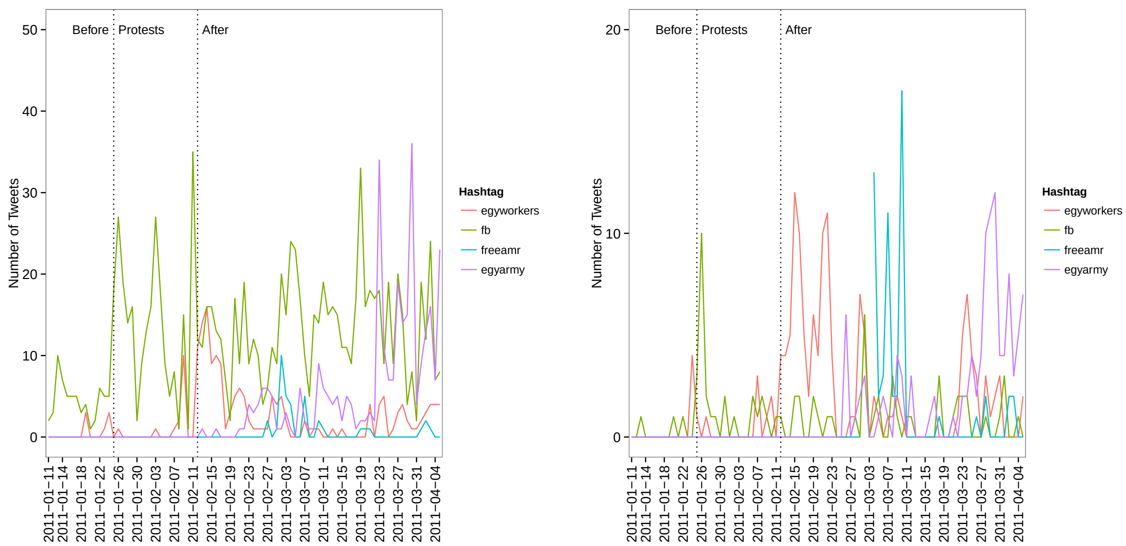
Figure 2.2: Egypt - Activists' and Non-Activists' Use of Common Hashtags



(a) Non-Activist Egyptians

(b) April 6^{th}

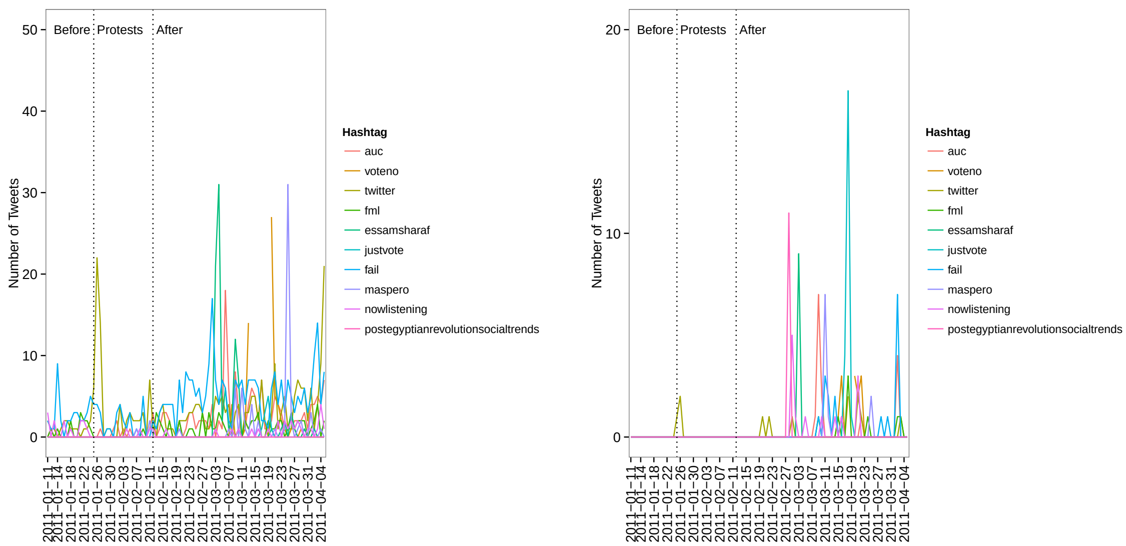
Figure 2.3: April 6^{th} - Activists' and Non-Activists' Use of Unique Hashtags



(a) Non-Activist Egyptians

(b) NoMilTrials

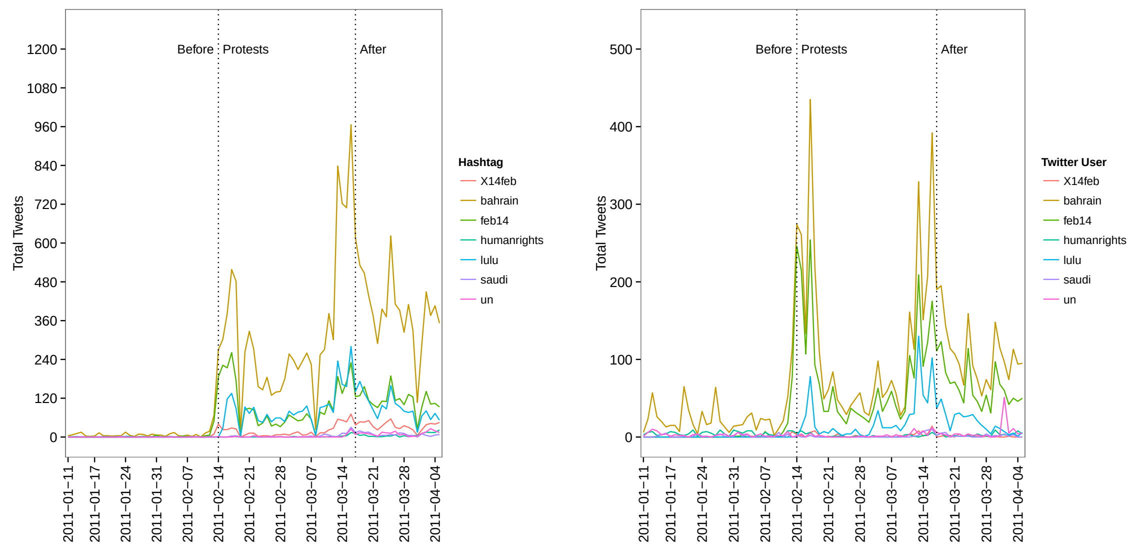
Figure 2.4: NoMilTrials - Activists' and Non-Activists' Use of Unique Hashtags



(a) Non-Activist Egyptians

(b) Anti-Sexual Harassment

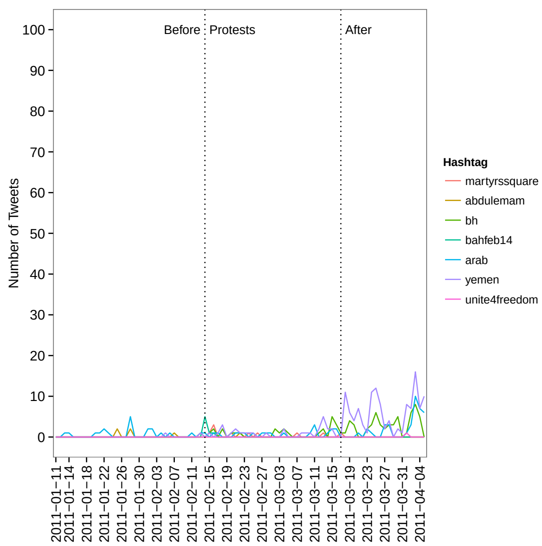
Figure 2.5: Anti-Sexual Harassment - Activists' and Non-Activists' Use of Unique Hashtags



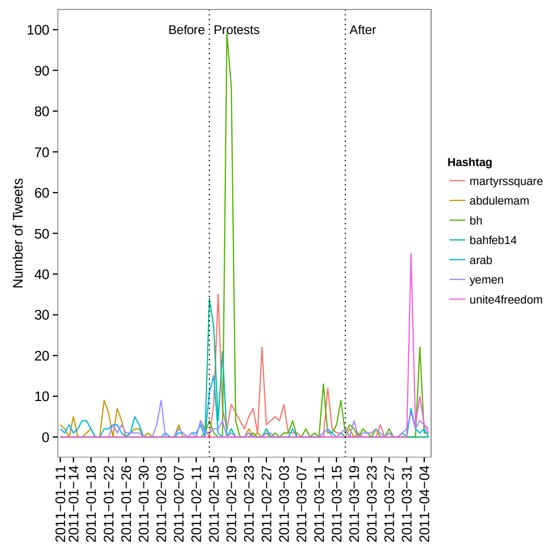
(a) Non-Activist Bahrainis

(b) Human Rights

Figure 2.6: Bahrain - Activists' and Non-Activists' Use of Common Hashtags



(a) Non-Activist Bahrainis



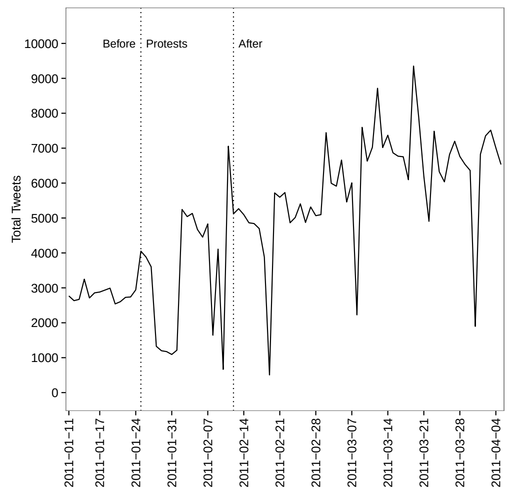
(b) Human Rights

Figure 2.7: Bahrain - Activists' and Non-Activists' Use of Unique Hashtags

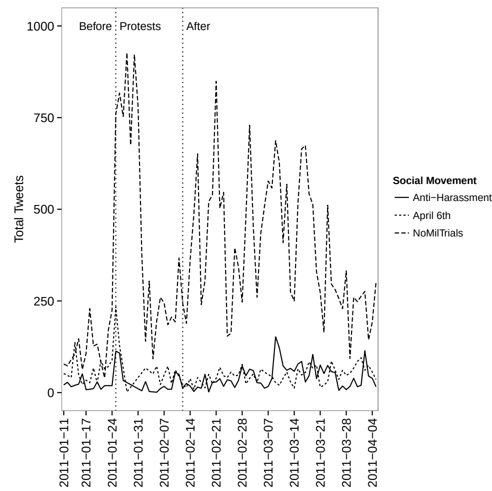
2.8.4 Blackout

It could be that a telecommunications blackout drive the (lack of) results. The Egyptian government blocked access to Facebook and Twitter starting on January 25^{th} [Ghonim, 2012, pg. 186], and on the evening of January 27^{th} it severed almost all telecommunications services, including the internet [Gunning & Baron, 2013, pg. 286]. Only one internet service provider, used by the Cairo stock exchange, government ministries, and some international hotels, remained operational. Just as protests crested in size, the number of tweets observed therefore dropped noticeably. This blackout, not the actual effect of activists, could therefore drive the lack of correlation observed between activists' behavior and protest.

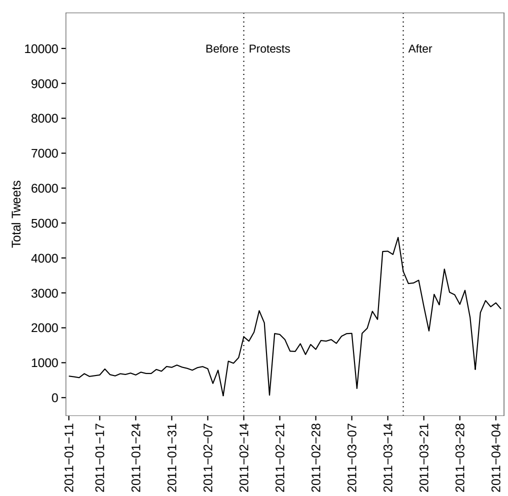
The blackout does not threaten inference. Though tweets were less frequent than before the blackout, they did not stop; reports confirm that activists and participants were able to use hotels' business centers and services like Speak to Tweet to produce and consume tweets [Khalil, 2011, pg. 200]. Figure 2.8 shows total tweet production for Egyptian and Bahraini activists as well as non-activists in Egypt and Bahrain. The daily count of tweets never reaches zero. Moreover, the variables measured are proportions, not counts. So long as the blackout does not change the kinds of activity that occur on Twitter, the statistics from that period will not differ because of the blackout.



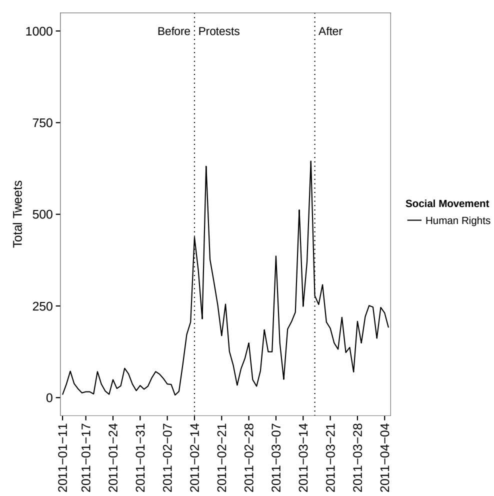
(a) Egyptian Non-activists



(b) Egyptian Activists



(c) Bahraini Non-activists



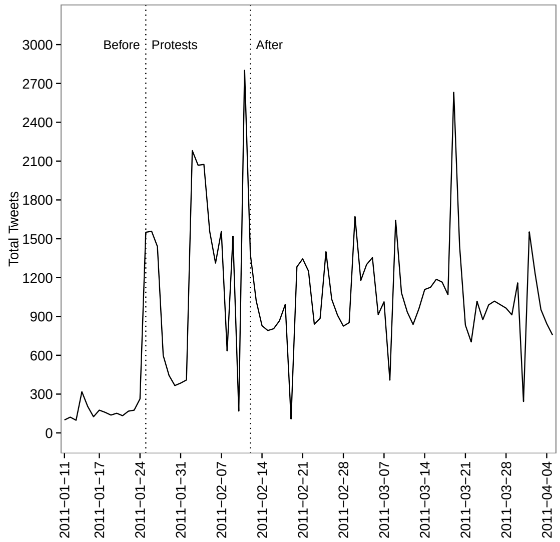
(d) Bahraini Activists

Figure 2.8: Activists' and Non-Activists' Differential Twitter Activity

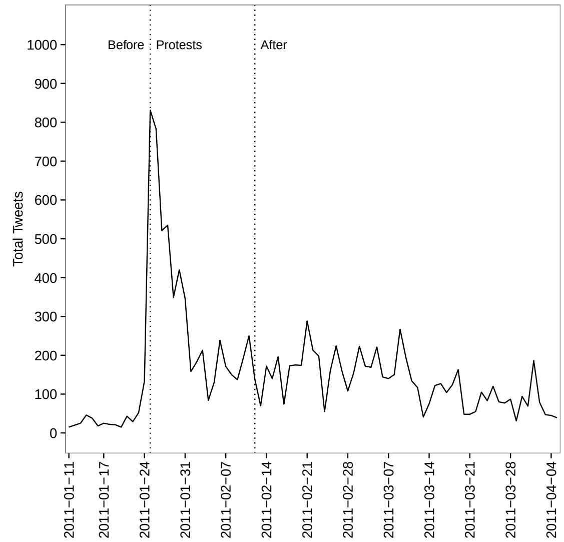
The lack of a result for activist coordination more likely results from a noticeable disconnect between topics activists and non-activists found salient during protests, as a comparison of hashtags used by activists and non-activists suggests. To reach this conclusion, the 25 most common hashtags of each movement were identified and separated into a set of those which all movements used and a set for each movement of hashtags used only within that movement. Those hashtags were then found in the Egypt and Bahrain data. If the frequency with which activists and non-activists used the sets of hashtags concords, the two groups most likely found the same topics salient; a difference suggests the topics activists found important were not considered so in the larger population.[14,15]

[14] Of the top 25 hashtags for each Egyptian movement, each movement uses "#dostor2011", "#egypt", "#jan25", "#libya", "#mubarak", "#syria", "#tahrir", and "#tunisia". In Bahrain, the common hashtags are "#14feb", "#bahrain", "#feb14", "#humanrights", "#lulu", "#saudi", and "#un".

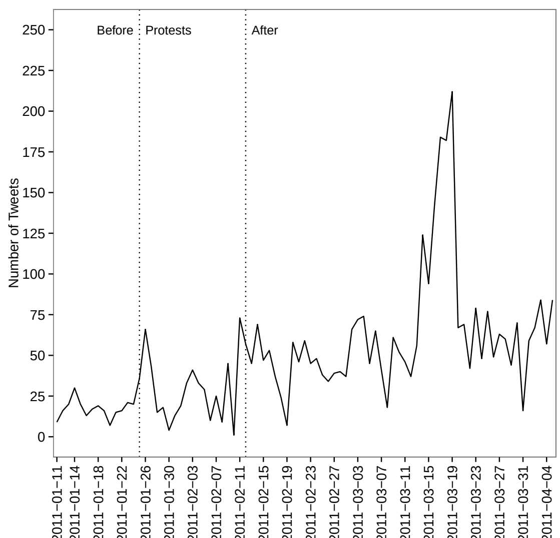
[15] The hashtags unique to April 6^{th} are "#revolution", "#tunisie", "#alarabiya", "#6april", "#tripoli", "#mobarak", and "#at"; "#mobarak" refers to the Hishram Mubarak Law Center, an organizing point for activists that government forces raided on February 2^{nd}, "#at" is from links to YouTube videos. Unique to the NoMilTrials movement are "#egyworkers", "#fb", "#freeamr", and "#egyarmy". "#auc", "#voteno", "#twitter", "#fml", "#essamsharaf", "#justvote", "#fail", "#maspero", "#nowlistening", and "#postegyptianrevolutionsocialtrends" are unique to Anti-Sexual Harassment activity. In Bahrain, the Human Rights movement is the only one to use "#martyrssquare", "#abdulemam", "#bh", "#bahfeb14", "#arab", "#yemen", and "#united4freedom".



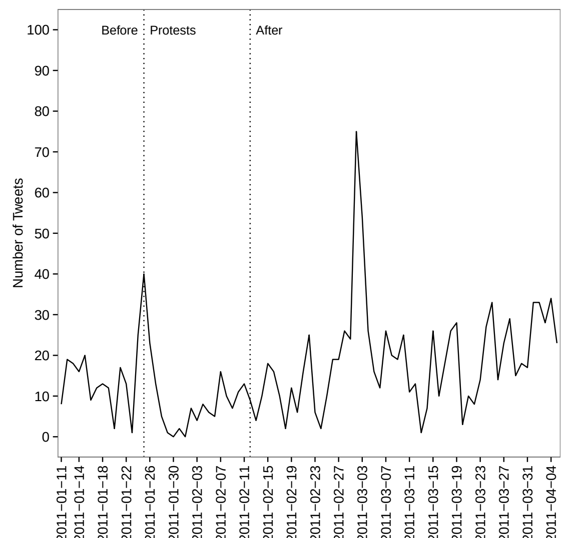
(a) Non-Activist Egyptians, Common #



(b) Activist Egyptians, Common #

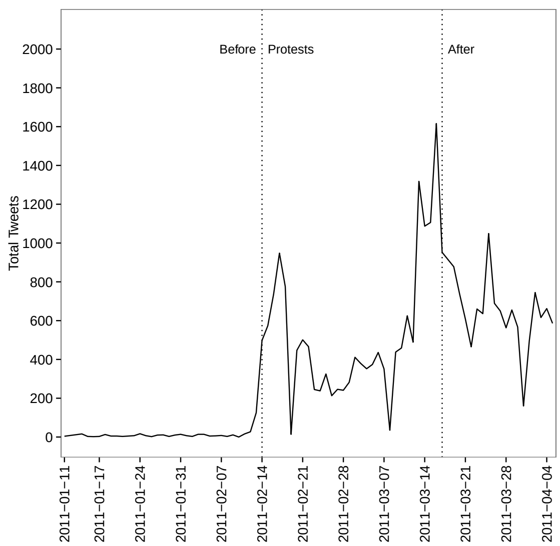


(c) Non-Activists, Activists' #

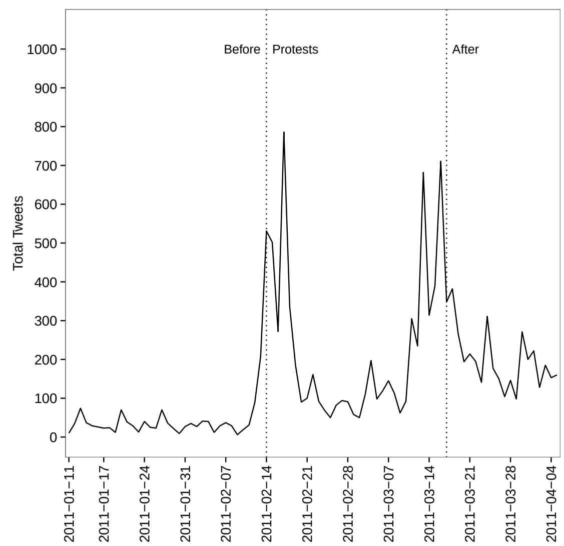


(d) Activists, Activists' #

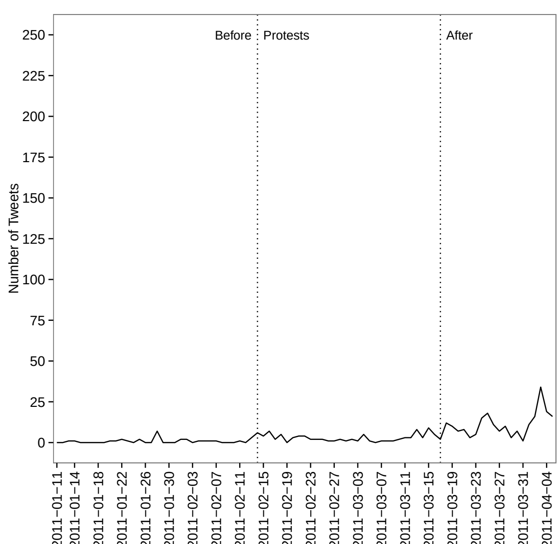
Figure 2.9: Egypt - Differential Hashtag Use



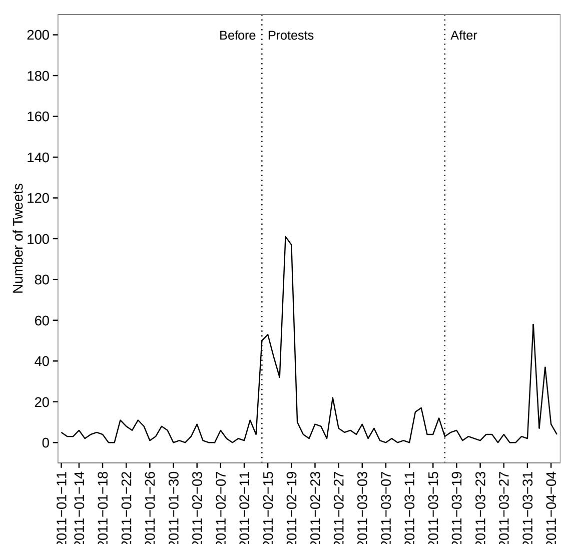
(a) Non-Activist Bahrainis, Common #



(b) Activist Bahrainis, Common #



(c) Non-Activists, Activists' #



(d) Activists, Activist #

Figure 2.10: Bahrain - Differential Hashtag Use

Figure 2.9a shows the prevalence of common hashtags amongst non-activist Egyptians; Figure 2.9b shows the same hashtags amongst activists. Three key patterns reveal themselves. First, both groups use the hashtags with the same changing frequency Figure 2.8 showed. Non-activists use the hashtags in increasing amounts during the 18 days, while activists cease using them about halfway through. Second, non-activists continue to use these hashtags after the protests, while activists continue to barely use them; by mid-March, activists rarely use the hashtags again. Third, there is more similarity between Figures 2.9c and 2.9d. Overall, the topics which are particular to activists are taken up by non-activists, but activists do not engage as much in broad topics that also resonate with non-activists.

Opposite trends appear in Bahrain, as Figure 2.10 reveals. Figures 2.10a and 2.10b show that the hashtags common to Bahrain movements are also common amongst non-activists. Yet the hashtags unique to activists are barely used amongst non-activists, as shown in Figures 2.10c and 2.10d. In Bahrain, there was thus accord on common hashtags but not on those unique only to the Human Rights movement.

A one-way street in hashtag topics appears to exist. Non-activists use hashtags specific to certain movements but also use general interest hashtags, whereas activists use their own hashtags at a much higher relative frequency than the general interest ones. Activists may not have influenced protests through coordination because they did not use hashtags prevalent amongst non-activists. There was thus a disconnect between the two groups during the protest period, and this disconnect affected the ability of activists to coordinate. Activists still affected protests in other ways, by increasing common knowledge and engaging in negotiation with the regime, but they did not affect them through hashtag coordination.

2.9 Acknowledgements

Chapter 2, in part, has been submitted for publication. I am the primary investigator and author of this paper. This chapter is made possible with support from the American people through the United States Agency for International Development (USAID) and under the Democracy Fellows and Grants (DFG) program implemented by IIE. The contents are the sole responsibility of the author and do not necessarily reflect the views of USAID, the United States Government, or IIE.

Chapter 3

Longitudinal Network Analysis with Incomplete Data

3.1 Abstract

Political scientists lack a low-cost methodology for analyzing structural properties of large scale networks. This paper shows how to analyze individuals' changing structural position at a daily level, using the social network Twitter. To do so, two innovations are introduced. First, one can infer when two individuals connect with other an arbitrary amount of time after they actually connected, a task made difficult by how Twitter delivers data to researchers. Communities which connect individuals from different countries can also be identified with this first method. Observing daily network change reveals changing communities and individuals' position therein. Second, a network measure from computer science, *neighbor cumulative indegree centrality (NCC),* is introduced; it preserves the rank ordering of individuals' centrality without the complete network data that those measures require. Combining the first method with the second creates daily data on network centrality. Moreover, these methods can be applied to a network after the period under study has past. Without these methods, daily data on the structural position of individuals would be prohibitively costly to obtain. These methods are demonstrated with 21 Twitter accounts from Bahrain and Egypt during a 3 month period in early 2011. Ground truth data on their number of followers confirms the accuracy of the *post hoc* inference, the activists' network centrality changes, both absolutely and relative to each other, and individuals who link activists in each country are identified.

3.2 Introduction

The increasing prevalence of digital communications technology - the internet and mobile phones - provides the possibility of analyzing human behavior in a level of detail previously unavailable. Blogs, content aggregation sites, internet fora, online social networks, and call data records provide text, social network, and location data that vary by the second. For political scientists interested in questions about elections, language, political communication, conflict, or spatial diffusion, among others, the rise of these technologies holds much promise [Bail, 2014].

While programming tools have lowered the barrier to acquiring, processing, and analyzing these data, it remains difficult to conduct structural network analysis. Structural network analysis requires complete network data, but complete network data is usually not available because of limits imposed by the data providers, and two existing approaches to circumvent this problem are less than ideal. First, researchers often use indegree centrality to measure network position of a node[1], but indegree centrality is less precise than measures that are calculable with complete network data because it does not take into account the position of that node's neighbors. Second, researchers can pool data sampled over a long time period, as connections between individuals should reveal themselves with a large enough sample, but pooling removes the possibility of observing the network change over time. This paper introduces two methods that allow for the analysis of network structure as it changes every day.

While digital communications technology provide data through numerous avenues, this paper focuses on one, the social network Twitter. With over 300 million

[1] "Node" means the entity that forms the network under study. In political science, a node will most often be an individual. For the rest of this paper, "node" means individual, individual means node.

accounts creating 500 million messages per day, it is one of the largest social networks. Its data are also relatively easy to access, compared to other platforms. While other social media platforms and websites, such as reddit or Instagram, also have easily accessible data, none are as general purpose as Twitter. Twitter's global reach, large user base, and data openness make it the preferred platform for large-scale studies of human behavior. Though Twitter is the preferred platform for analyzes of social networks, it makes such analyzes difficult because of how it provides data to researchers. Data provided as a streaming sample make structure difficult to see, while Twitter limits how often one can download data on connections between individuals. This paper introduces two methods that work within Twitter's limits to undertake longitudinal structural network analysis.

Section 3.3 introduces the 2 methods. The first method allows the researcher to infer when two individuals connect with each other, information not automatically provided. The second method introduces the neighbor cumulative indegree centrality (NCC) measure; NCC provides the same rank ordering as centrality measures that require complete network data, but it only requires a one-step breadth first search. Section 3.4 explains how to gather network data from Twitter and the specific data used here - a combination of data from the streaming and representational state transfer application programming interfaces (REST API) concerning 21 accounts in Egypt and Bahrain. Section 3.5 provides background on the substantive application of the new measures: activism during the Arab Spring in Bahrain and Egypt. 21 accounts in Egypt and Bahrain were identified by reading newspaper articles, NGO reports, and academic works [Gunning & Baron, 2013]. All of their tweets from January 11th, 2011 - April 5th, 2011 (58,376 total) were purchased from Sifter, a

3rd party reseller of tweets. This period encompasses time before, during, and after the main protests in Egypt and Bahrain. These data provide a ground truth against which to compare the estimated connection dates obtained using the first method in Section 3.3.

Section 3.6 presents the main results. The estimated connection dates from the first method are shown to accurately measure when follower connections actually formed, using the complete history of tweets from the 21 accounts. It is then possible to observe each country's network as it grows, revealing structural patterns across the accounts; eventually, the two national networks draw closer, as international communities of accounts follow accounts in each country. The second part of the section shows results using the neighbor cumulative indegree centrality measure. Section 3.7 concludes.

3.3 Two Measures for Network Analysis

3.3.1 Inference of Edge Formation

Twitter does not reveal when two users enter a relationship, so a researcher only knows that a connection exists but not when it formed.[2] Two pieces of information from the REST API ameliorate this situation.[3] First, the list of followers (or friends) that Twitter provides is sorted in reverse chronological order, meaning one can infer connection formations of the followers relative to each other.[4] Second, the

[2] "Following" is the fundamental building block of Twitter. When this paper refers to two accounts in a relationship or a connection existing between two accounts, it means that one follows the other.

[3] Section 3.4 explains Twitter and the REST API in detail.

[4] Twitter is an asymmetric network, and the terms "follower" and "friend" have different meanings. A "follower" is an account which has indicated to Twitter that it wants to automatically be made aware of the tweets of an account it follows, while a "friend" is the account being followed. If

REST API reveals when an account was created. These two pieces of information make it possible to accurately reconstruct when connections are formed.

When a relationship formed is estimated using the information Twitter provides through the GET users/ids endpoint of the REST API. This information includes the users screen name, self-reported location, preferred language, number of friends, number of followers, and, most importantly, date the user joined Twitter. The earliest an account could follow another is the day that account joined Twitter.

Using the date followers join Twitter allows for the approximation of connection formation. For each follower in a user's follower list, the earliest that follower could have formed a connection is the most recent Twitter joining date of all followers below that follower. Similarly, the latest that follower could have formed a connection is the first Twitter joining date greater than that follower's Twitter joining date for the followers above that follower in the follower list. These two dates, the earliest and latest a connection could have formed, bound the slicing procedure when subsetting the user's followers list.

An example, illustrated in Table 3.1, clarifies. Suppose User 1 has followers A, B, C, D, E, F, G, H, and I, with A the newest follower and I the oldest. Follower A joined Twitter on New Year's Day 2010 but could not have followed User 1 before 01.01.2013 because that is the most recent Twitter joining date of the 9 followers. Follower C joined at the same time as A but could have started following User 1 as early as 01.01.2012 because the latest any of Followers C through I joined Twitter was that day. Follower G's earliest possible connection date is the same as the day it joined Twitter because neither of the 2 already existing followers joined Twitter

B follows A, B is A's follower while A is B's friend.

Table 3.1: Inferring Follower Relationship Formation

Follower	Date Joined Twitter	Earliest Date Connection Forms	Latest Date Connection Forms
A	01.01.2010	01.01.2013	Day of API Request
B	01.01.2013	01.01.2013	Day of API Request
C	01.01.2010	01.01.2012	01.01.2013
D	01.01.2012	01.01.2012	01.01.2013
E	01.01.2009	03.01.2011	01.01.2012
F	03.01.2011	03.01.2011	01.01.2012
G	01.25.2011	01.25.2011	03.01.2011
H	01.01.2008	01.01.2009	01.25.2011
I	01.01.2009	01.01.2009	01.25.2011

after Follower G. Follower A's latest possible connection date is whatever day the follower list was downloaded, since Follower A is the newest follower of User 1; the same is true of Follower B because no subsequent follower (which is only Follower A) joined Twitter after Follower B. We can infer that Follower I connected to User 1 no later than 01.25.2011 because the first follower who connected with User 1 and had a Twitter joining date later than Follower I, Follower G, joined Twitter on 01.25.2011. Since Follower I is User 1's first follower and joined Twitter on 01.01.2009, a connection between User 1 and Follower I formed sometime during this range. The same is true for Follower H; Follower H could not have connected to User 1 before Follower I, even though Follower H joined Twitter a year earlier, because Follower H is closer to the top of the follower list. A researcher interested in the network of User 1 on 01.26.2011, the day after Egypt's first major protests, knows definitively that User 1 had two followers, H and I, that day but may have had 3, depending on when Follower G connected to User 1. Figure 3.1 provides pseudocode for an algorithm that assigns these bounds.

Data: List of accounts, Γ, to study

Result: Estimated date each follower ω_i of account γ_i started following γ_i

Select account γ_i

Download Ω_i from Twitter, the reverse chronological sorting of γ_i's followers

for *each* ω_i **do**

    Obtain from Twitter the date ω_i joined

end

for *each* ω_i **do**

    Find each follower who started following γ_i before ω_i

    Of these, find the most recent date

    Assign this date as the earliest possible date for connection between γ_i and γ_i

end

for *each* ω_i **do**

    Find each follower who started following γ_i after ω_i

    Of these, find the earliest date greater than when γ_i joined Twitter

    Assign this date as the latest possible date for connection between γ_i and γ_i

end

Figure 3.1: Pseudocode for Inferring Connection Date

Meeder et al. (2011) provide an analytic explanation of this process. This paper extends that work in three ways. First, it provides a method for estimating the upper bound of the follower connection date formation. Having a lower and upper bound for follower connection dates allows for more precise estimation of connection formation, though the bounds approach each other as the number of followers increases. Second, Meeder et al. (2011) works with celebrity accounts because they rapidly gain followers; the accounts in this sample show that this technique extends beyond celebrities. Third, the results show that measuring true changes in followers is most accurate when combining the streaming and REST APIs, whereas Meeder et al. (2011) use only the REST API.

3.3.2 Centrality Measurement with Incomplete Data

In network analyzes, *centrality* refers to a set of statistics which attempt to measure which nodes are most important. There are three main classes of centrality: betweenness, closeness, and degree-based.

A node with a high betweenness centrality connects many nodes of a network; using this measure, the most important node is that which is on the most paths connecting any two nodes. Closeness centrality refers to the mean distance between one node and all other nodes; using this measure, the most important node is that which has the shortest average distance between itself and all other nodes.[5]

The most common measures of importance focus on the number of connections a node has to other nodes. The sum of these connections gives the degree of a node, and a node with higher degree is assumed to have more importance than one with lower degree. Measuring only the sum of connections of a node is called degree cen-

[5] Distance is measure in the number of steps between two nodes. If A and B are connected to each other, their distance is 1. If A is connected to C through B, the distance AC is 2 while the distance BC is 1.

trality or, in a directed network, indegree or outdegree centrality. Degree centrality is appealing because of its simplicity, but it does not give an indication of a node's position in the larger network: a node may have high degree centrality, but if those with which it is connected have few connections, the node probably is not very important. Many measures therefore take into account the importance of a node's neighbors to calculate a node's centrality, the idea being that an important node has neighbors that are also important. There are various ways to calculate these measures, some of the most common being eigenvector centrality, Katz centrality, and PageRank; see Newman (2010 for a mathematical explanation of these measures. K-core centrality, another common measure, is explained shortly.

Eigenvector, Katz, PageRank, and k-core centrality have intuitive appeal, but they require having data on every connection in a network. Though large networks routinely contain many fewer connections than their theoretical maximum, the number of connections will still grow exponentially with the size of the network. Calculating these centrality measures is therefore difficult with large networks. Given this difficulty, degree centrality is the most common measure of centrality in large scale studies, especially those using Twitter [Kwak et al., 2010, Garcia-Herranz et al., 2014].

Across a wide range of observed social networks, the most influential users, those whose information diffuses furthest along a network, are located in the k-core [Pei et al., 2014]. (A k-core is a subset of nodes in the network in which all nodes have at least degree k.) As with betweenness, closeness, eigenvector, Katz, and PageRank centrality, finding the k-core requires complete network data and is therefore not feasible for analysis of large social networks. Pei et. al (2014) introduce a measure that preserves the rank ordering of node influence identified through k-core decomposition and outperforms degree and PageRank centrality. The measure is created by summing the indegree centrality of a node's neighbors, and the resulting sum, the number of followers the followers have, preserves the rank ordering of other centrality measures. I call this measure the *neighbor cumulative indegree centrality*, or NCC. Formally:

$$NCC_i = \sum_{1}^{j} d_j \tag{3.1}$$

For each node i, the neighbor cumulative indegree centrality is the sum of the indegree centrality d_j for each neighbor j. An illustration of NCC is presented in Figure 3.2.

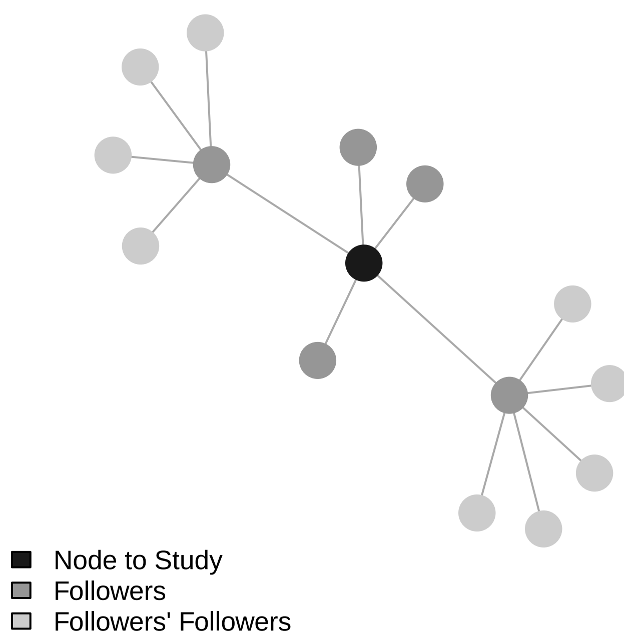


Figure 3.2: Degree Centrality = 5; NCC = 9

Intuitively, NCC works because it captures information on nodes up to 2 degrees away from the node for which NCC is calculated. The other centrality measures (except for degree centrality) operate recursively, meaning they capture information on a node's 3rd, 4th, 5th, ... nth connections as well. While the contribution to importance of a node's 3rd to n-th degree connections may matter, they should have less of an effect than a node's immediate and 2nd degree connections; empirically, this is the case [Christakis & Fowler, 2012]. Since these other measures operate recursively, they require complete network data; since NCC only requires data on the 2nd degree connections, it is must easier to obtain.

The key benefit of NCC is that it recovers other centrality measures that require complete network data. A comparison of sample strategies on four different networks finds that each sampling procedure requires a large network sample (over 50% of all nodes) before that sample's network characteristics converge to the full network's value [Lee, Kim & Jeong, 2006]. Sampling high-centrality nodes, performing breadth-first search out to 1 degree, and computing degree, betweenness, closeness, and PageRank centrality on the resulting subgraph preserves the rank ordering of node influence in addition to closely approximating those nodes centrality measures true values in the full network [Pei et al., 2014]. One does not know which nodes are highly central ahead of time, however, and collecting the amount of data required for Pei et. al (2014) is not feasible under most research budgets. Neighbor cumulative indegree centrality, in contrast, only requires a two-degree breadth-first search, significantly lowering the barriers to moving beyond indegree centrality. Section 3.4 explains how to use Twitter to calculate NCC.

While neighbor cumulative indegree centrality captures rank ordering that would be obtained using measures that require complete network data, it may still be preferable to have information on more than first-degree connections. In practice, such information is beyond reach. Because the number of connections in a network expands exponentially while Twitters rate limits are fixed, computing time increases supralinearly. For the 21 users in this study, their 1,908,134 followers have 506,821,726 followers; at 60 requests per hour returning a maximum of 5,000 followers per request, one connection would need just over 70 days to download the list of 2nd-degree followers. Assuming 45% of those are unique (the percentage from the crawl of followers for this paper), one computer would require almost 132 days to download data on each unique user. While this number is probably an overestimate since some of the 2nd-degree followers may have been followers of one of the 21 accounts, the rate at which the download time increases as a function of degrees from a seed node is unknown. A complete crawl of Twitter conducted in July 2012 supplements 2 machines that could make 20,000 requests per hour and 2 that could make 100,000 with 550 machines using the normal rate limits; this crawl required 4 months and 4 days [Gabielkov, Rao & Legout, 2014]. The 4 machines with higher rate limits were whitelisted, a now defunct practice by which Twitter gave certain machines preferential access to their data. A similar crawl without whitelisted machines would therefore take about double the time according to those authors' estimates.

3.4 Network Data

3.4.1 Twitter

Social scientists first started to analyze data based on the internet in the early 2000s. The link structure of the web itself earned early attention [Broder et al., 2000], as did instant messaging services [Cameron & Webster, 2005, Li, Chau & Lou, 2005, Leskovec & Horvitz, 2007]. As Web 2.0 technologies - websites where users provide content - rose from the ashes of the dot-com bust, MySpace [Hinduja & Patchin, 2008, Thelwall, 2008, Tufekci, 2008] and blog platforms [Herring et al., 2004, Radsch, 2008] received more attention. Online fora [Nielsen, 2012] and content aggregation sites are also common objects of study [Lerman & Ghosh, 2010, Gilbert, 2013]. Facebook was founded in 2004, Twitter in 2006; as their user base has grown, so has the number of academics studying them. For a review of the literature about Facebook, see Wilson (2012); no parallel exists for Twitter.

Five features of Twitter have driven its popularity for academic study. First, it is one of the largest social networks, with 284 million active users in almost every country and over $1 billion of annual revenue [Twitter, 2013, 2014]. These users include heads of state, companies, non-profit organizations, international non-governmental organizations, celebrities, athletes, journalists, academics, and, primarily, everyday people. In the United States, as of December 2013, 18% of adults use Twitter; men and women use it equally, users are largely between 18-39 years old, and roughly an equal percentage of adults from different education and income brackets use it. 46% of its users use it daily and 21% at least once a week (versus 63% and 22% for Facebook) [Duggan & Smith, 2013]. Twitter therefore provides a cross-section of almost any group in which a researcher would find interest. The only social networks larger are Facebook, WeChat, and Sina Weibo.

Second, users produce large quantities of data, 500 million messages per day. All these people and messages mean that Twitter mirrors vast segments of the population that would otherwise require large teams of researchers, and very large budgets, to analyze concurrently. Taken together, these first two characteristics mean that almost any event is recorded on Twitter, and many events are predictable with it. Twitter has been used to predict stock market changes [Bollen, Mao & Zeng, 2011], box office returns [Asur & Huberman, 2010], coups [Kallus, 2013], emergent memes [Garcia-Herranz et al., 2014], crime [Gerber, 2014], and winners of professional football games [Sinha et al., 2013].

Third, it makes these data relatively easy to obtain. Twitter makes its users' data available through two application programming interfaces (API), the streaming API and Representational State Transfer (REST) API, that are accessible to anyone with a Twitter account.[6] The streaming API provides as much as 1% of all tweets every day and accepts filters based on keywords, user location, user IDs, or language. Through the REST API, one can download specific tweets, 3,200 of a user's most recent tweets, a list of who a user follows or who follows the user, and user profile information. One needs some programming knowledge to interact with the APIs, though there are a large number of software libraries to access Twitter, including in Python and R.

Fourth, Twitter is an excellent data source for network and non-network analysis. Since the service is explicitly structured as a network - connections between accounts are the fundamental building blocks of the user experience - researchers interested in diffusion processes and emergent behavior find Twitter a natural source. Twitter, because its 1% stream delivers tweets without information on the tweet author's social network, is also a compelling source for researchers interested in polling and event prediction [Gayo-Avello, 2013]. Non-network applications of Twitter are

[6] Before the switch to v1.1 of the APIs, Twitter allowed third parties to provide interfaces that allowed individuals with no programming experience to access its APIs. Now, that capability no longer exists, raising the barrier to entry for acquiring data.

discussed further in the conclusion.

Sixth, Twitter has a norm of public conversation that does not exist on many social networks, such as Facebook. While Facebook also provides an API, most users choose not to make their information publicly available. To gain access to a user's information, one needs to design a Facebook app that the user installs or work with Facebook's research team. This team maintains veto power over research proposals and publications, and the recent controversy over manipulation of Facebook feeds has caused the company to tighten control over its research team [Kramer, Guillory & Hancock, 2014].

3.4.2 Network Data from Twitter

Networks can be reconstructed from either API, though Twitter inhibits the extraction of network data. The most common avenue for acquiring Twitter data is by connecting to its streaming API, a real-time sample of what is being said, when, and, with a little more work, where. Each tweet delivered through the streaming API also records the number of followers and friends of the author of the tweet; these measures are equivalent to indegree and out-degree centrality, respectively. But the streaming API does not automatically provide more complex measures of centrality. Because the REST API provides follower and friend lists, only it permits the reconstruction of all of a user's connections. A researcher interested in the complete social network of an individual therefore has to use the REST API.

The main downside of the REST API is rate limits, restrictions on how frequently one can request data from Twitter. For example, one can only download the follower or friend information for a maximum of 15 accounts every 15 minutes. Since each request returns up to 5,000 followers or friends, a user with more than 5,000 of either will require multiple requests, and each request counts towards the 15 per 15 minute limit. Since the number of friends or followers grows exponentially as one crawls a network and a researcher is most likely interested in networks of hundreds or thousands of people, the rate limits quickly become become a barrier.

Different sampling strategies decrease the time required to crawl Twitter. First, the task can be split across multiple computers. Because downloading the information on each user does not depend on who has been downloaded - there are no interconnections within the follower or friend lists - the task can be assigned to as many computers as possible. Splitting the task across multiple computers is the easiest way to quicken the download process [Gabielkov, Rao & Legout, 2014]. Second, one can prune any individual with too many followers. This will bias the resulting network away from the core, an acceptable outcome conditional on the research question. Third, the researcher can accept only the first 5,000 followers or friends of an account. Twitter returns friends and followers in ordered lists of 5,000, where newer connections are returned before older ones. Since each returned list counts towards the rate limit, an account with 75,000 followers therefore takes as long to download as 15 accounts with 5,000 each. In practice, the distribution of followers and friends follows a power law [Kwak et al., 2010], so most downloading time is concentrated on a subset of accounts with many followers. Restricting a download to only one page of followers will therefore provide a substantial speed increase. Fourth, one can download 2nd-degree followers or friends based on a filtering criterion. For example, if the researcher only cares about English speaking accounts, it may make sense to only download followers whose account language is English. This process assumes metadata on the followers is already downloaded, so this strategy would not save time for either of the methods used in this paper.

The most common, and fastest, approach to obtaining network data is to use the streaming API. Using the streaming data, the researcher can create network connections based on retweets or user mentions. For example, “RT @ZacharyST: Hello world” implies some connection between @ZacharyST and the account retweeting. Similarly, in “@ZacharyST: LeBron James is over-rated”, a connection can be inferred between @ZacharyST and the tweet author who mentions @ZacharyST. Using retweets to infer connections will overweight popular users, as they get retweeted the most [Boyd, Golder & Lotan, 2010, Kwak et al., 2010, Suh et al., 2010], while user mentions will produce the fewest edges [Conover et al., 2011, Zamal, Liu & Ruths, 2012*b*]. Using the retweet or user mention network allows researchers to create network data from data not designed for that purpose, and it reduces the amount of computing time required to start network analysis.

Two problems arise with retweet and user mention networks. First, the ties do not necessarily reflect actual Twitter connections, as an account can retweet or mention another account without following that account. If a researcher is interested in social connections between individuals, then it is preferable to observe those actual connections, not communication behavior that may not reflect actual social relationships. Second, creating these networks through the streaming API makes temporal analysis difficult. The streamed data are a sample; measuring daily change in a mention or retweet network would therefore reveal a new network each day. Analysis of networks from the streaming API therefore pool data collected over longer time periods, foreclosing the possibility of longitudinal analysis. Retweet and mention networks are therefore best suited for questions about those networks; they have more difficulty answering questions involving social structure or temporal change. Whether or not creating daily mention or retweet networks changes the structural properties of each network is an open question.

3.5 Activism During the Arab Spring

3.5.1 Background

The phrase “Arab Spring” refers to the counter-regime protests that occurred across the Middle East and North Africa starting at the end of 2010. The spark was the self-immolation of Mohammed Bouazizi, a fruit vendor in the provincial town of Sidi Bouzid, Tunisia, on December 17th, 2010. An act born of economic and political frustration, the burning led to protests across the country, culminating in President Zine al-Abidine Ben Ali’s flight from power on January 14th, 2011. Copycat immolations throughout the region arose, and the next major protests started in Cairo on January 25th, with President Hosni Mubarak surrendering power on February 11th. Almost every Arab country experienced some level of protest, with most governments responding with a mixture of repression as well as political and economic concessions. The civil wars in Libya, Yemen, and Syria are direct results of each country’s Arab Spring protests, and Tunisia is the only country that enjoys a nascent representative political system.

Egypt and Bahrain featured some of the most contentious mobilization and repression. Egypt experienced 18 days of major protest during which state security forces engaged in sustained violence, mass arrests, and the regime made token political concessions; the armed forces appear to have forced President Mubarak from power, and protests against the subsequent period of military rule continued through 2011. 2012 presidential elections led to the ascension of Mohammed Morsi, the Muslim Brotherhood's candidate, but an Army coup on July 3rd, 2013 ended the democratic period. The armed forces now clearly control the country, and a low-level insurgency has commenced. Bahrain, home of 570,000 citizens, had 150,000 protesters at its peak, the highest level of mobilization during the Arab Spring; protests, which cut across sectarian divisions, decreased after military intervention from the Gulf Cooperation Council. Low-level anti-regime action continues outside of Manama, the capital.

From Egypt and Bahrain, 42 activists representing 5 social movements were identified, 19 of whom were active on Twitter prior to each country's first protests. In Egypt, activists from the April 6th youth movement, the No Military Trials campaign, and the Anti-Sexual Harassment movement were chosen; in Bahrain, the human rights community and February 14th youth coalition were chosen, though only the human rights community was active on Twitter before the start of protests. The final 19 activists represent the 3 social movements in Egypt and Bahrain's human rights community. 2 Bahraini government accounts were also identified and collected, raising the final count of accounts under analysis to 21.

3.5.2 Acquiring Activism Data

There are 3 methods for acquiring data from Twitter: purchase, download in real time via the streaming API, or download old tweets and account metadata through the REST API. Longitudinal network analysis can be done using only the REST API, but this paper also purchased tweets from a reseller in order to verify results.

Twitter's REST API provides an account's 3,200 most recent tweets and metadata on the account; for longitudinal network analysis, the important pieces of information are when the account joined Twitter and the account's followers. Each of the 21 accounts was submitted to different endpoints of the REST API, leading to up 3,200 tweets per account and the list of followers of each count. Each follower in the list of followers was then submitted to the REST API in order to learn when that follower joined Twitter and how many followers the follower had. This information is then sufficient to measure daily changes in network structure, as explained in Section 3.3.

While the data acquired through the REST API suffice to undertake longitudinal network analysis, data from Sifter, a 3rd-party tweet reseller, were used to verify the results. Sifter allows one to purchase old tweets based on various parameters, one of which is account name; the price of the request is a function of the number of days and tweets in the request. Sifter does not return the entire tweet object, but it does include pertinent information such as the tweets text, authors username and user ID, a timestamp, the number of followers, and number of friends. The metadata on number of followers can then be used to slice the list of followers for an account, providing more precise cutpoints than the inference introduced in Section 3.3. Section 3.6 uses these cutpoints to verify the accuracy of the inference of the date a connection forms between an account and a follower.

Purchasing old tweets is preferable for two reasons. First, tweets delivered through the REST API provide the number of followers of the account of interest on the day the tweets are downloaded, not the day the tweets were authored. Second, it allows the researcher to analyze events that happened before the research question was developed. In other words, there is no need to purchase historic tweets if one was able to monitor the accounts under study in real time, but it is rare that a researcher will have the prescience to follow these accounts for an indefinite amount of time, waiting for an unknown event to occur. For example, this study could have started following the 21 accounts in late 2010 or early 2011, but doing so would have required predicting the Arab Spring and which actors would be important during those events. Because tweets through the REST API do not provide accurate follower counts, even downloading accounts' historic tweets as soon as an event of interest occurs will not facilitate network analysis.

3.6 Application

3.6.1 Reconstructing Daily Network Change

This section demonstrates how the procedure in Section 3.3.1 accurately measures the true number of followers and reveals changing network structure. There are two ways to measure a users change in followers over time: either observe that user in real-time (with the streaming API) while frequently downloading their followers' list (via the REST API) or estimate, later and indirectly, that change. The former is most precise but requires the researcher knows which accounts he or she is interested in before the date range over which an account is observed for a study. Estimating the change indirectly, through the REST API, is therefore how most longitudinal analyzes will proceed. This section demonstrates that the latter accurately recovers the results from the former and can show daily change in network structure.

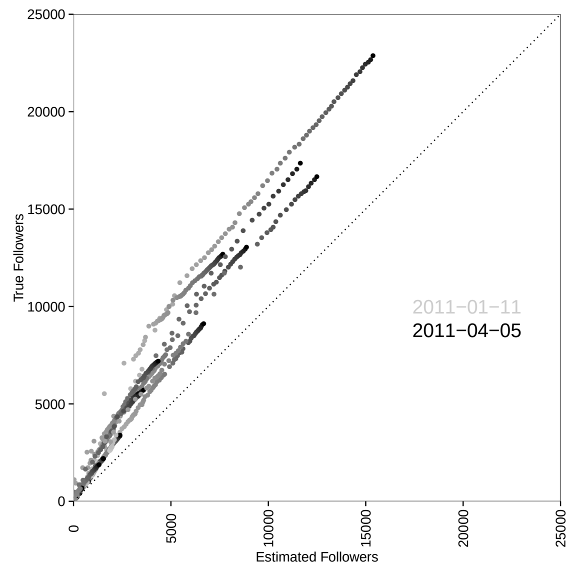
When one receives a tweet from Twitter, the tweet indicates how many followers its author has. For example, if a tweet from @Mrmeit on January 24th, 2011 indicates he has 200 followers, the 200 followers from his follower list provides his social network at that time. If @Mrmeit has 5 new followers the next time he appears in the dataset, the 5 newest followers are 201st-205th in the list. Following this procedure, one can reconstruct the evolution of a social network, though with two caveats. First, a user who stops following @Mrmeit will no longer be on the follower list, so there will be some imprecision at the cutoff points.[7] Second, the time resolution of changes in the network is limited by how often a user is observed, and an account that tweets infrequently will appear less often in the streaming sample. Since streamed data are a sample, one likely does not have every tweet from a particular user and maybe not even one tweet per day per user. For example, User A may have 2 tweets 5 days apart in a streamed sample; if User A gained 10 followers between two tweets, it is unclear when during those 5 days those connections formed.

The tweets purchased from Sifter are used to verify the reconstructed follower network. Since they are a complete history, they do not suffer from the second problem. But the followers list for each account was downloaded in the summer of 2014, so there will still be noise if some follower relationships ended.

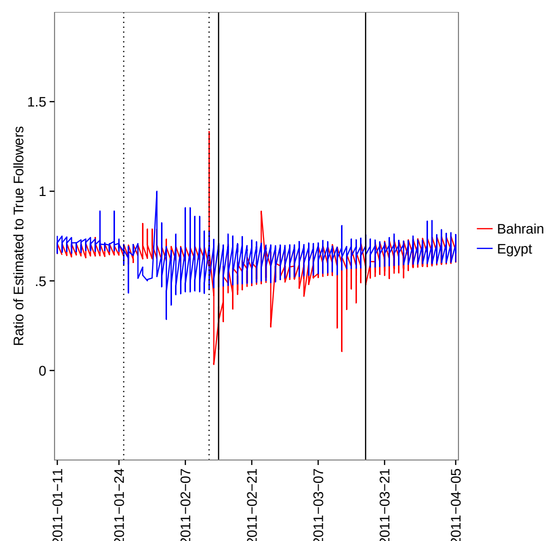
Figure 3.4a shows that the *post hoc* estimated number of followers, calculated as described in Figure 3.1 for each day from January 11th, 2011 - April 5th, 2011, linearly predicts the true number of followers. The true number comes from observing

[7]Having the follower list for each time @Mrmeit appears in the dataset obviates this problem, as the lists could be compared to each other to see who has stopped following him. This approach is impractical, however, for 2 reasons. First, it relies on knowing ahead of time in which users the researcher is interested so that the follower list can be downloaded each time the account appears in the dataset. Second, Twitter's rate limits mean that scaling this approach would require significant engineering.

the change in number of followers based on the Sifter tweets. Because users can stop following an account, the followers list was downloaded after the period of study, and Twitter removes users from the followers list once they stop following an account, the estimated number of followers underpredicts the true number.



(a) Approximation of Followers' Daily Change



(b) Accuracy by Country, Day

Figure 3.3: Verification Against Ground Truth Data

Accounts are shaded from light to dark gray based on how close to April 5th, 2011 they are. The estimated number of followers explains 98.09% of the variance in the number of true followers, with half of the remaining variance explained by group fixed effects; both these estimates are based on a linear model not shown here.

The residual increases as a function of the estimated number of followers, but this heteroskedasticity is constant as a percentage of an account's followers.

Figure 3.4b shows that the estimated number of followers is usually 67.53% of the true number of followers. This relationship whether or not the results are pooled by country; aggregating observations by group does not change the trends. The dashed lines correspond to the start and end of protests in Egypt, the solid in Bahrain. The *post hoc* measure performs less consistently, though does not appear biased, during these protest periods, suggesting that the measure may perform less well when the number of followers fluctuates rapidly. Overall, the *post hoc* measure of followers consistently approximates the true measure, suggesting it can be used when the true number of friends is not observable.

Figures 3.5 and 3.6 show network structure in Egypt and Bahrain on January 24th, 2011 and April 4th, 2011, respectively.[8] These dates represent both countries before the start of protests and after, providing a high-level view of how network structure changed over the course of three tumultuous months. Each node represents a group of users who follow the same accounts and whose Twitter profile uses the same language. Language is assumed to proxy for location, though English is highly indeterminate and is often the default language used when Twitter is not available in a users preferred language. Nodes are colored by language and sized by the number of accounts in the node. Nodes for the 21 accounts are labeled and contain only the 1 account. The unlabeled nodes are referred to as communities when discussing results. Edges between communities are omitted to enhance legibility.

---

[8]A Fruchterman-Reingold algorithm generates each layout, with equal weighting assigned to all edges. The Fructherman-Reingold is a spring-based algorithm: it attempts to maintain equal distance between all nodes. In the resulting graphs, nodes closer together are most similar based on the communities which follow them.

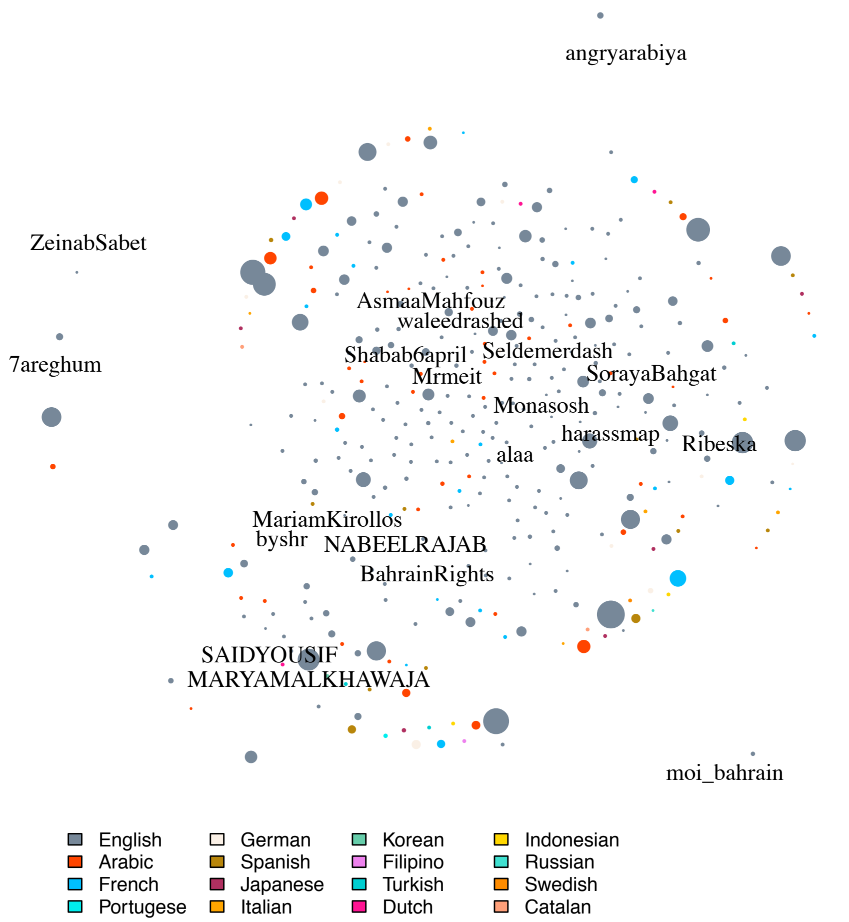


Figure 3.4: Bahrain and Egypt Activists, 01.24.2011

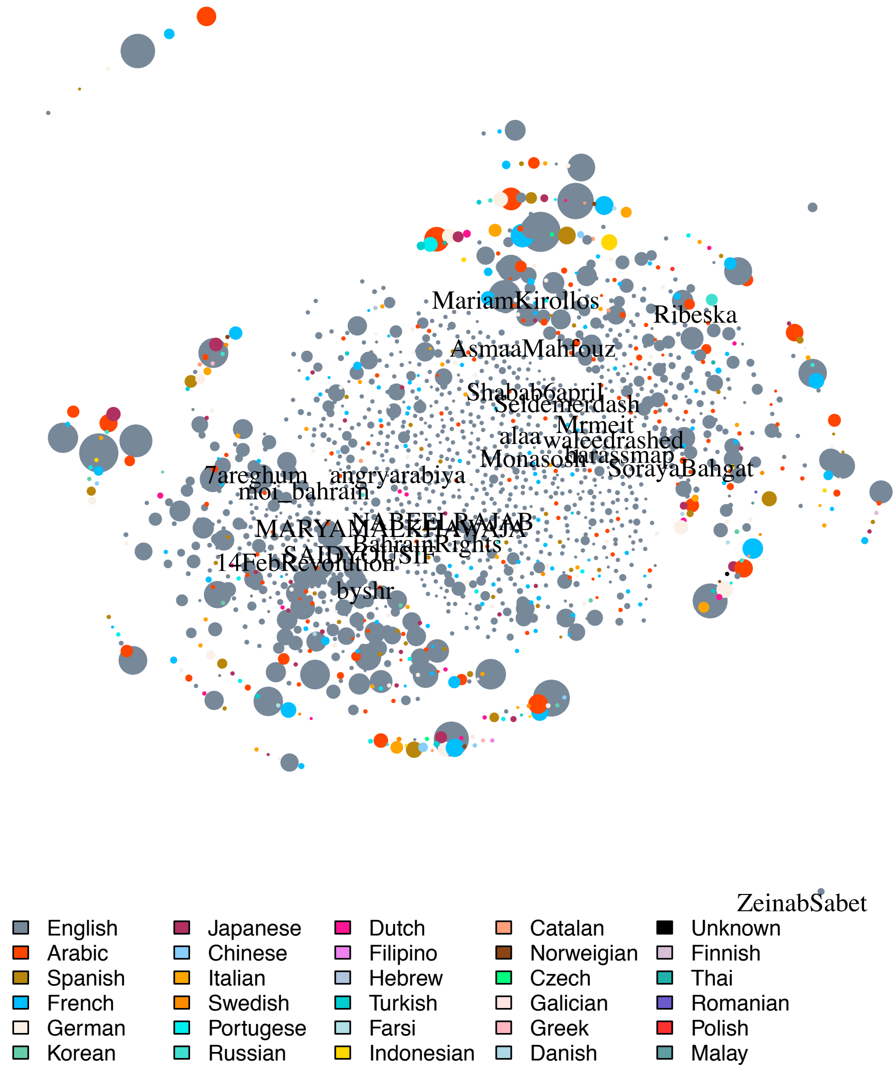


Figure 3.5: Bahrain and Egypt Activists, 04.04.2011

Figure 3.5 reveals three main structures. First are the 3 isolated seed accounts, those whose followers do not follow any other accounts. Only one of these, @ZeinabSabet, is from Egypt, though it has the most followers of the 3. The 2 Bahraini accounts are the government or affiliated with it. @moi_bahrain is the Ministry of the Interior, and @7areghum is an account of unknown origin that tweeted photos, names, and addresses of protestors [Bassiouni et al., 2011]; they are the only 2 in this sample related to the government. Note that @angryarabiya is almost an isolate, with only two communities following her; one follows only her, another follows @Monasosh.

There exists a 2nd cluster of accounts from Egypt and a 3rd of accounts from Bahrain. The Egyptian cluster is in the center of the graph, the Bahraini one in the lower-left quadrant. The exception is @angrayarabiya, who is closer to the Egyptian accounts because very few people follow her, so those that do that also follow accounts in Egypt bring her closer to the Egypt cluster. The least central accounts in Egypt are individuals associated with the Anti-Sexual Harassment movement: @Ribeska, @SorayaBahgat, and @Seldemerdash, in addition to @ZeinabSabet. Finally, @BahrainRights and @NABEELRAJAB are closer to the Egyptian network than @SAIDYOUSIF and @MARYAMALKHAWAJA because more of their followers also follow Egyptian accounts. Accounts were not labeled ahead of time as being from Egypt or Bahrain; the clustering by country derives from the network of their followers.

Figure 3.5 also reveals the less developed structure of the Bahraini social network. Internet penetration in Bahrain was, and is, higher than in Egypt [*International Telecommunications Union Statistics*, 2014], but its prevalence did not translate into a more connected political social media sphere for the accounts in this study. While determining the impact of social media on contentious action is a burgeoning research field [Howard et al., 2011, Gunning & Baron, 2013], the fact that Egypt and Bahrain both experienced high levels of turnout but had very different online network structures may suggest a small role for social media [Zeitzoff, Kelly & Lotan, 2015].

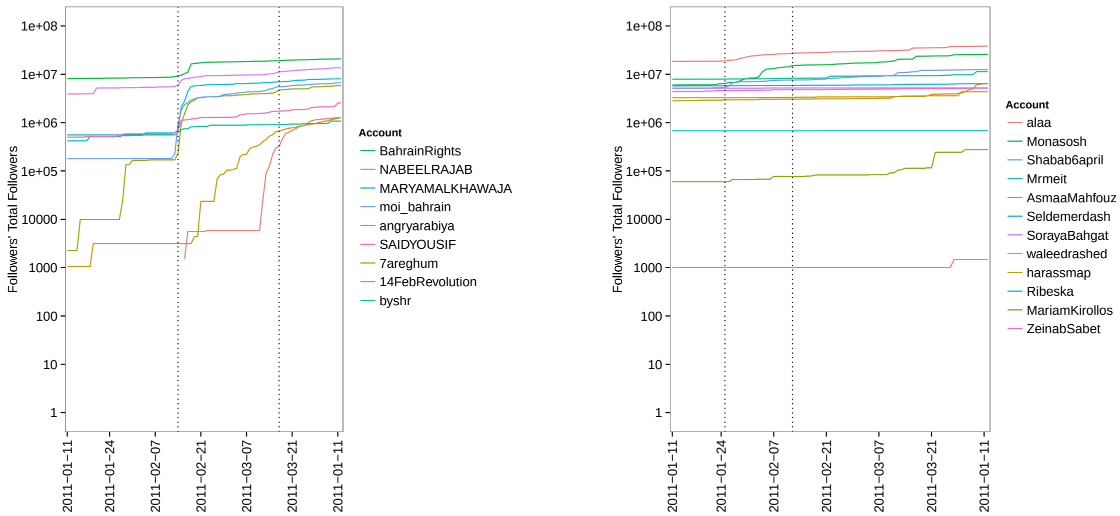
By April, 2 weeks after Bahrain forcefully suppressed protests, its network looked very similar to Egypts, as seen in Figure 3.6. The pro-Bahrain accounts are now very similar to the other Bahrain accounts, though they have many groups of followers that only follow them. There is also still separation between @Bahrain-Rights and @NABEELRAJAB from the other Bahrain accounts, and they are joined by @angryarabiya and @MARYAMALKHAWAJA, the head of the Bahrain Centre for Human Rights since Nabeel Rajabs arrest. These accounts are set off from the other Bahraini accounts due to their followers that also follow Egyptian accounts. @Monasosh and @Seldemerdash are the Egyptians with the most followers who also follow Bahraini accounts, just as before the protests. The communities that link accounts in Egypt and Bahrain are much smaller in size than those that only follow accounts in either country. Only the followers of one account, @ZeinabSabet, do not also follow other seed users.

Comparing Figure 3.5 and Figure 3.6, the overarching trend is one of growth and densification. There are many more communities in early April (Figure 3.6) than the end of January (Figure 3.5) for Bahrain and Egypt in addition to a similar increase in the number of communities that follow accounts in both countries. The communities also have more members on average at the end of the study, though the border-spanning communities are smaller than the intranational ones. Finally, the 21 seed accounts have a much more international audience than before: 25 unique languages on April 4th compared to 16 10 weeks prior.

3.6.2 Daily Changes in Influence

To measure neighbor cumulative indegree centrality, the user ID of each of the 21 seed account's followers was downloaded from Twitter's GET users/ids endpoint, returning 4,229,373 results containing 1,908,134 unique followers. Each user ID was then submitted to Twitter's GET users/lookup endpoint, providing data such as when the user joined Twitter, their self-reported location, their default language, and how many tweets they have authored. These first-degree followers themselves have 506,821,726 followers. Since it would take 6 months to download the 2nd-degree connections, and weeks more to download metadata for each ID, data on 2nd degree connections were left alone.

Figure 3.7 presents the change in neighbor cumulative indegree centrality over three months in Bahrain and Egypt. The first vertical line represents the start of protests, the second the end. Each country's legend is ordered from highest to lowest values of NCC at the end of the period.



(a) Bahrain (b) Egypt

Figure 3.6: Reconstructed Temporal Change in Influence

A few results emerge from Figure 3.7. In both countries, relative influence is stable: the rank ordering of NCC on January 11th, 2011 looks very similar to that on April 5th, 2011. Even though every account except for @Ribeska gains NCC, very few gain influence at a quicker rate than their peers. In Bahrain, a notable change is @angryarabiya, who moves from 2nd least influential to 5th most; that account belongs to the daughter of Nabeel Rajab (@NABEELRAJAB), a human rights advocate who led - he is now imprisoned - the Bahrain Center for Human Rights (@BahrainRights). @moi_bahrain, the Ministry of the Interior's account, is the 4th most influential at the end of the study, an increase of 2 spots. @byshr, the account of the Bahrain Youth Society for Human Rights, experiences the steepest decline, moving from 3rd to last. Egypt's relative ordering is more stable. @Shabab6april experiences the greatest change in NCC, moving from 5th to 3rd. @monasosh experiences a large increase in absolute influence, but she only moves from the 3rd to 2nd most influential account.

Both countries' accounts also experience the greatest changes in influence around their protest periods. Each country's users start to gain influence days before the start of protests. Most continue to gain influence during the protest period, and some stabilize after while others continue to gain in importance.

Finally, comparing the NCC across Bahrain and Egypt reveals intriguing network properties. The Bahrain accounts start and end with lower average influence than the Egyptian ones. Egypt, on the other hand, has higher variance in influence, though the result may be due to the accounts sampled. The three least influential Egyptians accounts, the relative ranking of which do not change, are accounts for individuals associated with the Anti-Sexual Harassment movement. That movement has been more peripheral to Egyptian politics than those sampled in Bahrain. Excluding those 3, the Egyptian accounts have greater influence and lower variance than the Bahraini ones. Why countries' networks have different structural properties is outside of this paper's scope but has started to receive some attention [Zeitzoff, Kelly & Lotan, 2015].

3.7 Conclusion

This paper has shown that researchers can now observe daily changes in network structure. While theoretically possible without online social networks, complexity and cost rendered such an undertaking effectively impossible until recently. Combining various components of Twitter's application programming interface, a researcher can infer with high precision when one user starts to follow another, and this information can be combined with a new measure of centrality to measure individual accounts' network position at a daily level. The global reach of Twitter, large quantity of public communication that occurs on it, and ability to obtain these data at low cost make it an ideal source for scholars interested in a broad range of questions, including networks.

The main barrier presently facing researchers are programming rate limits. Future work should explore how to approximate neighbor indegree centrality without having to sample all of a node's followers. Because of the way Twitter returns data, the approximation would need to work with the newest followers of an account. This paper has also only treated one direction of an asymmetric network, treating accounts as emitters of information. But individuals also consume information, and the consumption network should change over time as well. The symmetric network - where each connection represents mutual following - will also reveal patterns about more intimate types of relationships. How these networks change over time remains an open question.

Though this paper has focused on network analysis using Twitter, the platform is also amenable to large scale text and geospatial analysis. Twitter is a particularly attractive data source for text analysis because barriers to entry are substantially lower than for appearing in other venues such as a news report or in Congress. People who are not likely to appear in those sources are therefore much more likely to appear on Twitter, casting light on swathes of previously unobservable behavior. While this outcome is true of the internet more broadly - barriers to entry are similarly low for joining most social networks, posting to web forums, or starting a blog - Twitter is unique in its combination of size and public communication. Elites who are likely to appear in traditional text sources usually have Twitter accounts (every Senator and 430 of the 435 members of the House have official Twitter accounts, as do most news organizations), meaning Twitter records communication behavior from all strata of society. No other data source exists which is simultaneously comprehensive and accessible.[9]

It is also possible to map Twitter activity to specific places, allowing scholars to connect patterns on Twitter to offline events. Location information comes from two sources, accounts choosing to provide their GPS coordinates or self-reporting their location as part of their profile. Tweets with GPS coordinates, which represent 2-3% of all tweets [Leetaru et al., 2013], provide the most precision in estimating location. They can be used to map language communities at the city and neighborhood level [Hale, Gaffney & Graham, 2011, Mocanu et al., 2013] and unemployment [Llorente et al., 2014]. They also reveal that offline geography conditions online network connections [Yardi & Boyd, 2010, Conover et al., 2013]. In addition to representing few tweets, they are more prevalent in urban areas and among higher-income users, so the extent to which they are reliable depends on the research question [Malik et al., 2015].[10]

[9] A growing body of work uses Twitter to understand mass behavior. Tweets can be used to measure sentiment across time and space in ways consistent with time of day and season [Dodds et al., 2011, Golder & Macy, 2011] as well as distance from home [Frank et al., 2013], and this sentiment can predict changes in the Dow Jones Industrial Average [Bollen, Mao & Zeng, 2011] and the Standard & Poor 500 [Zheludev, Smith & Aste, 2014]. Textual analysis also suggests that members of Congress' rhetoric follows that of those they represent, not the other way around [Barberá et al., 2014]. Tweets can also reveal dynamic patterns of political polarization in countries as diverse as the United States [Barberá, Jost, Nagler, Tucker & Bonneau, 2015] and Egypt [Borge-Holthoefer et al., 2015]. Text in tweets can also be used to predict upcoming crimes [Gerber, 2014] and is used as one of many textual inputs to predict political unrest [Kallus, 2013]. Studies of Twitter during local emergencies suggest that information is shared in a two-step flow model [Katz, 1957], with most people sharing information they first obtained from news organizations [Starbird & Palen, 2010, Yardi & Boyd, 2010].

[10] To increase the number of geolocated tweets in a sample, scholars infer location based on a user's self-reported location in their profile, though the user does not have to identify their location. This location can be any text string, so some users report their location as "Space" or "Under the Sea", but most reported locations correspond to actual places. Including users' self-reported location, anywhere from 34% [Leetaru et al., 2013] to 50% [Conover et al., 2013] to 66% [Hecht et al., 2011] of tweets have location information, the vast majority of which are at the city level. It is possible to assign location information to accounts based on the context of their tweets [Cheng,

This paper has focused on networks in Bahrain and Egypt, but the methods could be applied in any locality to any group of people. 21 accounts were chosen because those were relevant to the behavior being studied, but dozens to hundreds of accounts could be analyzed before the researcher needs to consider adding hardware. These methods could provide insight into how network structure affects the diffusion of political ideas, sentiment, or mobilization. They can also be used to identify international brokers, those individuals who connect communities across national boundaries [Carpenter, 2011].

Finally, these methods can be used to study offline networks. It is common for studies of networks and political behavior to run surveys and ask respondents to name their friends [**?**Opp & Gern, 1993]. Modifying this approach, a researcher could ask those the respondent names how many friends they have or even ask the respondent how many friends she or he thinks each of the friends has. This information would be enough to generate NCC scores for the original respondents. Generating the NCC from offline data allows researchers who do not use social network datasets or who are interested in samples of individuals not on these networks to also approximate centrality when full network data are not available.

Caverlee & Lee, 2010, Stefanidis, Crooks & Radzikowski, 2011]. It is unknown whether accounts with self-reported location differ from those without, so the benefit of this added parsing needs to weighed against the computational cost. The most global source of location is an account's time zone, as Twitter automatically assigns a time zone to each tweet based on the location where the tweet author registered [Lotan et al., 2011].

3.8 Acknowledgements

Chapter 3, in part, has been submitted for publication. I am the primary investigator and author of this paper. This chapter is made possible with support from the American people through the United States Agency for International Development (USAID) and under the Democracy Fellows and Grants (DFG) program implemented by IIE. The contents are the sole responsibility of the author and do not necessarily reflect the views of USAID, the United States Government, or IIE.

Chapter 4

Twitter as Data

4.1 Abstract

The rise of the internet and mobile telecommunications has created the possibility of using large datasets to understand behavior at unprecedented levels of temporal and geographic resolution. Online social networks attract the most users, though users of these new technologies provide their data through multiple sources, e.g. call detail records, blog posts, web forums, and content aggregation sites. These data allow scholars to adjudicate between competing theories as well as develop new ones, much as the microscope facilitated the development of the germ theory of disease. Of those networks, Twitter presents an ideal combination of size, international reach, and data accessibility that make it the preferred platform in academic studies. Acquiring, cleaning, and analyzing these data, however, require new tools and processes. This paper introduces these methods to social scientists and provides scripts for downloading Twitter data.

4.2 Introduction

The increasing prevalence of digital communications technology - the internet and mobile phones - provides the possibility of analyzing human behavior in a level of detail previously unimaginable. Blogs, content aggregation sites, internet fora, online social networks, and call data records provide text, social network, and location data that vary by the second. For political scientists interested in questions about elections, language, political communication, conflict, or spatial diffusion, among others, the rise of these technologies holds much promise [Grimmer & Stewart, 2013, Bail, 2014].

These data require new tools to acquire, process, and, sometimes, analyze.

These tools are no more difficult to learn and use than other qualitative and quantitative methods, but they are not commonly taught to social scientists. There also exists no canonical text, in journal or book form, that explains the strengths and weaknesses of these data and tools. This paper provides a systematic introduction to these data sources and the tools needed to benefit from them.

While digital communications technology provide data through numerous avenues, this paper focuses on one, the social network Twitter. With over 280 million accounts creating 500 million messages per day, it is one of the largest social networks. Its data are also relatively easy to access, unlike Facebook's. While other social media platforms and websites also facilitate data access, none are as general purpose as Twitter. Twitter's global reach, large user base, and data openness make it the preferred platform for large-scale studies of human behavior.

This paper explains how to use Twitter to undertake large-scale studies of human behavior. Section 4.3 provides a brief literature review of other digital communications technologies and explains advantages of Twitter in more detail. Section 4.4 discusses the three methods for acquiring Twitter data (purchasing, collaboration, or downloading), hurdles to downloading data, and programming tools. Section 4.5 discusses the types of substantive issues for which Twitter data are best suited. Section 4.6 concludes with a discussion on the implications of large-scale data analysis for the social sciences.

4.3 Why Twitter?

Scholars first started to analyze data based on the internet in the early 2000s. The link structure of the web itself earned early attention [Broder et al., 2000], as did instant messaging services [Cameron & Webster, 2005, Li, Chau & Lou, 2005, Leskovec & Horvitz, 2007]. As Web 2.0 technologies - websites where users provide content - rose from the ashes of the dot-com bust, MySpace [Hinduja & Patchin, 2008, Thelwall, 2008, Tufekci, 2008] and blog platforms [Herring et al., 2004, Radsch, 2008] received more attention. Online fora [Nielsen, 2012] and content aggregation sites are also common objects of study [Lerman & Ghosh, 2010, Gilbert, 2013]. Facebook was founded in 2004, Twitter in 2006; as their user base has grown, so has the number of academics studying them. For a review of the literature about Facebook, see Wilson (2012); no parallel exists for Twitter.

Six features of Twitter have driven its popularity for academic study. First, it is one of the largest social networks, with 284 million active users in almost every country and over $1 billion of annual revenue [Twitter, 2013, 2014]. These users include heads of state, companies, non-profit organizations, international non-governmental organizations, celebrities, athletes, journalists, academics, and, primarily, normal people. In the United States, as of December 2013, 18% of adults use Twitter; men and women use it equally, users are largely between 18-39 years old, and roughly an equal percentage of adults from different education and income brackets use it. 46% of its users use it daily and 21% at least once a week (versus 63% and 22% for Facebook) [Duggan & Smith, 2013]. Twitter therefore provides a cross-section of almost any group in which a researcher would find interest. The only social networks larger are Facebook and Sina Weibo.

Second, Twitter produces a lot of data, 500 million messages per day. All these people and messages mean that Twitter mirrors vast segments of the population that would otherwise require large teams of researchers to analyze concurrently. Taken together, these first two characteristics mean that almost any event is recorded on Twitter, and many events are predictable with it. Twitter has been used to predict stock market changes [Bollen, Mao & Zeng, 2011], box office returns [Asur & Huberman, 2010], coups [Kallus, 2013], emergent memes [Garcia-Herranz et al., 2014], crime [Gerber, 2014], and winners of professional football games [Sinha et al., 2013].

Third, it makes these data relatively easy to obtain. Twitter makes its users' data through two application programming interfaces (API), the streaming API and Representational State Transfer (REST) API, that are accessible to anyone with a Twitter account. Before the switch to v1.1 of the APIs, Twitter allowed third parties to provide interfaces that allowed individuals with no programming experience to access its APIs. Now, that capability no longer exists, raising the barrier to entry for acquiring data. Nonetheless, because of Twitter's popularity, there are a large number of software libraries to access Twitter, including in Python and R. One needs some programming knowledge to interact with the APIs, though not as much as just a few years ago.

Fourth, the APIs make it easy to tailor the data one receives to a specific research question. One can receive as much as 1% of all tweets every day (from the streaming API) or filter the tweets received based on keywords, user location, user IDs, or language used. Through the REST API, one can download specific tweets, 3,200 of a user's most recent tweets, a list of who a user follows or who follows the user, and user profile information. In other words, though Twitter has pioneered many "big data" technologies, one does not necessarily need to possess these skills to access Twitter's data. One may need to learn new skills to gather the data, but modeling and visualizing those data can be done with old tools of the trade.

Fifth, Twitter is an excellent data source for network and non-network analysis. Since the service explicitly structured as a network - connections between accounts are the fundamental building blocks of the user experience - researchers interested in diffusion processes and emergent behavior find Twitter a natural source. But Twitter, because its 1% stream delivers tweets without information on the tweet author's social network, is also a compelling source for researchers interested in polling and event prediction [Gayo-Avello, 2013].

Sixth, Twitter has a norm of public conversation that does not exist on Facebook. While Facebook also provides an API, most users choose not to make their information publicly available. To gain access to a user's information, one needs to design a Facebook app that the user installs or work with Facebook's research team. This team maintains veto power over research proposals and publications, and the recent controversy over manipulation of Facebook feeds has caused Facebook to tighten control over its research team [Kramer, Guillory & Hancock, 2014].

4.4 Acquiring Data from Twitter

There are 3 approaches to acquiring data: purchase, collaboration, or using Twitter's application programming interface (API). This section details each in turn.

4.4.1 Purchase

If one is interested in tweets from the past, the most thorough approach is to purchase them from a vendor. (Section 4.4.3 describes how to download some old tweets for free.) Companies which provide access to old tweets pay a large, undisclosed licensing fee to Twitter, and their main customers are marketing and public relations firms. Since many companies provide this service, this section focuses on four of the most prominent.

The main vendor is GNIP; it was founded in 2008 and started licensing Twitter's data in 2010, and Twitter purchased the company outright in 2014. While GNIP's target market is businesses, anyone can purchase their old Twitter data using their Historical PowerTrack application programming interface (API). GNIP claims that prices start at $500, but a project will more likely spend upwards of $5,000 purchasing data. The price is a function of the number of tweets that will be returned and the timespan of the request, though the final price requires consultation with a sales representative. GNIP provides its own metadata as well, including expanded links, a Klout score, language detection, and enhanced geo-information. Though GNIP provides a programming interface, a one-time purchase is best handled through contacting a salesperson directly. GNIP has offered Data Grants, free downloads of tweets to winning applicants; its first, and so far only, competition saw 1,300 entrants compete for 6 grants.

DataSift is a third major reseller of archived tweets. They provide the same services as GNIP and Topsy, and, like those services, provide a programming interface that uses their own syntax to filter historic data.[1] DataSift's historic data starts on January 1st, 2010, whereas GNIP and TOPSY have every tweet since 2006. DataSift charges $1 per 5 hours of computation time, plus $.10 per 1,000 tweets a query returns. To estimate the cost of a dataset, the researcher has to submit a query to DataSift's API. DataSift will return the estimated number of matching items, which, when divided by 10, will give the cost in dollars for the tweets themselves; the

[1] DataSift, GNIP, and Topsy also ingest other datasources such as Wikipedia, Reddit, YouTube, and WordPress.

final cost will also include computation time. Two features distinguish DataSift from its competition. First, it accesses more services' data than other services. Second, DataSift has also published a Python library to access its API. An annual subscription costs $200,000 - $300,000.

The most cost effective method to acquire historic tweets is through Texifter's Sifter tool. Texifter is a 3rd party vendor that interfaces with GNIP. The main advantage of Sifter is that it provides an online interface to GNIP's API so that a user can test a variety of search parameters. Sifter will then estimate the number of tweets the download would contain and provide a cost estimate. Sifter also appears to offer a lower price point than GNIP, though whether that is true or not is unclear since GNIP's pricing structure is opaque. Sifter then transfers the data to Texifter's DiscoverText tool, an online interface through which a coding team can search, filter, code, or analyze the returned tweets. The tweets can then be downloaded, though Twitter's terms of service require that Sifter only allow 50,000 tweets per day to be downloaded, regardless of how many match a user's query.

4.4.2 Collaborate

On April 14^{th}, 2010, Twitter and the Library of Congress announced the Twitter Research Access project, a collaboration to make every tweet ever published available to researchers [Stone, 2010]. Scheduled for completion in 2013, the project still has not resulted in an available archive. Updates from the Library have been intermittent, though it is clear it has at least all tweets from 2006, when Twitter started, through the end of 2010. Disconcertingly, a report from the Library in 2013 suggested that the hardware necessary to enable fast searches of the archive are "cost-prohibitive and impractical for a public institution" [*Update on the Twitter Archive At the Library of Congress*, 2013]. The most detailed report on the project and its current status is to be found in Zimmer (2015).

@Greptweet hosts greptweet.com, a website that returns a user's 3,200 most recent tweets. This website is the most user friendly method to get a user's tweet history, but it suffers two main disadvantages. First, it only returns the tweet ID, tweet timestamp, and tweet text; while one can retrieve the full tweet based on the ID (see Section 4.4.3), it would be preferable to have the data without programming. Second, one can only get tweets for one user at a time, which is a much slower process than submitting requests through the API. greptweet.com is therefore best for retrieving data for a small group of people who do not tweet frequently.

The Twitter Collection and Analysis Toolkit (TCAT) allows an interested party to connect to Twitter's streaming API, follow users or keywords, and create network graphs through Gephi, all without any programming expertise [Ward, 2014]. TCAT is best for identifying accounts which are influential in talking about a particular product. The service is no longer open to the public, though inquiries are accepted on an *ad hoc* basis [Groshek, 2015].

CrisisNet is part of Ushahidi, an open source digital platform that supports data-gathering from the internet and mobile phones, with a focus on the developing world and public sector implementations. CrisisNet ingests data from Ushahidi, Facebook, Twitter, Instagram, and ReliefWeb, processes that content into a standard form, and makes each post from each source available through an application programming interface. It is not a complete archive of content from the data sources it monitors; instead, it pays attention to specific accounts that CrisisNet users identify.

Once an account is identified, its data from that point forward is part of CrisisNet. The platform is still a small operation and has the steepest learning curve of these tools.

Finally, many academics have already collected, or collect continuously, their own Twitter data, though Twitter's terms of service prevent the mass sharing of tweets. Collaborating with someone who already has Twitter data on a topic or timeframe of interest is often the quickest, most inexpensive method of acquiring tweets.

4.4.3 Acquire on Your Own

The third option for collecting tweets is to collect them oneself. Advantages of this approach include being able to define search terms, not relying on others for data, and, depending on how much data is involved, cost. Disadvantages include a steeper learning curve than purchasing or working with others, difficulty accessing historic data, and needing to maintain one's own infrastructure. Section 4.4.3 and 4.4.3 explain the two application programming interfaces (APIs) for acquiring data for free and what kinds of data are available from each.[2]

REST API

The Representational State Transfer (REST) API provides access to past tweets, user data, and social structure.

User's Tweets Twitter allows anyone to download an account's previous 3,200 tweets through the GET statuses/user_timeline endpoint. This endpoint accepts 180

[2] "Free" in that Twitter does not charge for these two methods. One still needs hardware with which to store and analyze the data.

requests per 15 minutes and returns up to 200 tweets per request. An account with 3,200 or more tweets will therefore require 16 requests. With each tweet, Twitter returns metadata on the tweet author, but that metadata reflects when the API request was made, not when the tweet was created. @Greptweet is an interface on top of this endpoint.

Specific Tweets One can download more than 3,200 of a user's tweets if the identification number of each tweet is known. Twitter's Terms of Service prevent researchers from sharing original tweets; an individual who has tweets, whether streamed or downloaded, can only share the tweet ID (or user ID if one wants to share users). A researcher who would like to replicate other work or use previous tweets in original research is therefore reliant on the goodwill of the original acquirer of the tweets and having the programmatic ability to download the tweets. Fortunately, Twitter's rate limits are generous for downloading tweets, and the code to do so is simple. One can download 18,000 tweets every 15 minutes (100 for each of 180 calls), equivalent to 1,728,000 per day. Downloading user information is subject to the same limits. The most difficult part of downloading old tweets or user information is finding the Twitter IDs. One is reliant on another researcher's generosity or finding pertinent IDs already posted online [Freelon, 2012].

If one would like to download more than 1.7 million tweets per day, there are two options. The easiest approach - split the list of IDs into small chunks and submit those chunks at the same time - will not work because Twitter only allows 1 connection per IP address. The two options for downloading tweets at a higher rate are the two options to get around the IP limit. The first approach, which violates the spirit of Twitter's limit, is to route the requests through proxy servers. The second option is to use multiple computers. It is not difficult to launch multiple virtual instances using a hosting solution such as Amazon Web Services; the main drawback with a hosted solution is cost. Otherwise, friends' computers or old machines are perfect for this sort of task.[3]

Search Tweets It is possible to query Twitter's GET search/tweets for old tweets matching certain parameters. This method is to be used cautiously, however, as Twitter returns only some tweets from the previous 6-9 days and is not clear on how it chooses which to return. Only 100 tweets per request are returned, up to 180 requests per 15 minutes. To avoid receiving the same 100 results per request, pass the lowest tweet ID of the returned tweets to the max_id parameter in the subsequent call, and repeat this process as much as necessary. In addition to search terms, Twitter allows for filters based on language and latitude and longitude pairs; one can also specify for Twitter to return the most recent or most popular matching tweets, or a mix. One study has found that results from GET search/tweets do not match those from the random sample [González-Bailón et al., 2012].

User's followers Reconstructing network connections is slightly more difficult. Two endpoints, GET followers/list and GET followers/ids, provide information about followers. The former provides fully hydrated user objects for each follower, up to 15 followers per 15 minutes. The latter provides only the identification number of followers, but it does so for 75,000 followers per 15 minutes; those numbers can then be fed to GET users/lookup, from which up to 18,000 completely hydrated user objects are returned every 15 minutes. GET followers/list therefore saves one step but is slower than using GET followers/ids with GET users/lookup.

[3]Old computers are perfect for any Twitter task where rate limits force the task to take a long time.

User's friends The same logic holds for retrieving who a user follows (that user's friends, in Twitter parlance). One connects to GET friends/ids instead of GET followers/ids, but those friend identification numbers are fed to GET users/lookup.

Twitter returns the follower and friend list in reverse chronological order but does not reveal when either connection is formed. Section **??** explains how to infer connection dates using the REST API.

Streaming API

This section details how to collect data in realtime via Twitter's streaming API.[4] There are two levels of access available, the 1% vs. 10 % stream. Twitter removed free access to the 10% stream - variously called the garden hose or fire hose - in early 2011, and one now has to apply and pay for access. Twitter does not disclose the price of connecting to the garden hose; anecdotally, it appears to require large grants for an academic to afford it. The 1% stream remains free and probably will be forever, as it is the connection developers use to build Twitter-related products.

Random sample By default, Twitter returns a random 1% sample of tweets as they are written. This sample comes out to about 5 million tweets and 12 gigabytes of raw data per day. Because each returned tweet object consists primarily of metadata, preprocessing or post processing the tweets can reduce storage requirements without reducing the amount of useful data. Twitter does not disclose how it chooses the sample, leading to concern about the representativeness of the stream versus the complete Twitterverse. Comparisons of the stream to the complete Twitterverse find the stream does not differ in a meaningful manner [Morstatter et al., 2013, Valkanas

[4] The streaming API technically has 3 endpoints: GET statuses/sample, GET user, and GET site. Academics will only need to work with GET statuses/sample, so that is the connection assumed for the rest of the paper.

et al., 2014].

One can also request Twitter filter the results from the streaming endpoint. If a filter is given, Twitter returns all tweets matching the request *up to 1% of all tweets*. For example, if one asks for every tweet with the keyword “LeBron” and tweets with that word constitute .1% of all tweets, Twitter will return every single tweet containing that word. This design is extremely advantageous for researchers, as it means the streaming sample can often become a streaming population. With filters, the streaming API can provide a researcher with every tweet of interest, though the researcher will have to know ahead of time what filters are of interest. To continue the keyword example, if the researcher connects to the stream without specifying “LeBron” only 1% of all tweets containing that word will be returned, 2 order of magnitude fewer tweets than requesting tweets specifically with “LeBron”.

Language sample When connecting to the streaming API, one can request only tweets in a certain language. Twitter will then return all tweets in that language up to the 1% ceiling.

Geographic sample The streaming API can also return tweets from within a box defined by two coordinate pairs. The bounding boxes are not used in conjunction with other filters. For example, asking for tweets from San Francisco and tweets in Spanish will return all tweets from San Francisco (regardless of language) and all tweets in Spanish. Since 2-3% of tweets contain GPS coordinates [Leetaru et al., 2013], passing the coordinate pairs [-180,90,180,90] - a box around the world - will return 33% to 50% of all tweet with GPS coordinates. Twitter accepts up to 25 bounding boxes per connection. The streaming API does not use a user’s self-reported location.

Specific Text Twitter will return tweets containing a user-supplied string, and multiple strings can be passed. This functionality can be used to search for specific hashtags, individual words, links (Twitter will search the expanded URL of a shortened link), retweets, or mentions of a user. 400 pieces of text can be passed per connection.

Specific People One can submit to the streaming API specific user identification numbers and receive all tweets the users create, all tweets the users retweet, replies to tweets of the users, and retweets of the users' tweets. 5,000 people can be followed per connection. This feature is especially useful when the accounts to be studied are known. The best way to identify accounts is through lists, curated collections of accounts other users have created. For example, Twitter maintains the "US Senate" list, a list of the Twitter accounts for each Senator (every Senator has a Twitter account). A researcher could ask the REST API for the user identification number of each member of this list and then pass those numbers to the streaming API. Every tweet from every Senator would henceforth be downloaded by the user, assuming the 100 Senators never account for more than 1% of all tweets.

Except for the GPS bounding boxes, Twitter's documentation is not clear on how these parameters interact with each other. More likely than not, they are additive, e.g. following a specific user and asking for tweets in Spanish will probably return tweets in Spanish and tweets from that user.

4.4.4 Access Constraints and Replication

While Twitter's streaming and REST APIs are powerful, they have 7 quirks to be aware of when writing code. These quirks render some kinds of analyzes, especially those relying on the REST API, more difficult than they otherwise would be. They also raise the costs of replication, though full replication remains possible.

First, an overarching restriction is that Twitter imposes limits on how quickly one can retrieve data from the REST API.[5] Requests are counted in 15 minute windows, and most API endpoints allow 15 requests per window. Each request, however, may return multiple matches, and the number of matches returned is not constant across endpoints. For example, Twitter returns 5,000 followers per request, so 75,000 followers can be downloaded in one window; when asking Twitter for metadata about an account, 100 accounts per request and 180 requests per 15 minutes are allowed, allowing for metadata on 72,000 accounts every 15 minutes.[6] If one's code needs to query an endpoint more than the rate limit allows, the code must itself; otherwise, Twitter will sever the connection and reserves the right to ban the account from querying the APIs again. The rate limits make it very difficult to reconstruct complete social networks, an issue returned to in Section **??**. More detail on rate limits is available at Twitter's developer website.

Second, Twitter only allows one connection per IP address to any part of its API. For example, downloading the 1% stream and a separate stream returning only tweets from the United States would require two separate IP addresses. Similarly if one wants to connect to the stream while parsing accounts' followers. Without this restriction, the rate limits would be meaningless. Nonetheless, the ability to create virtual machines on demand, using a product like Amazon Web Services, makes the IP address restriction less onerous than it otherwise would be.

[5] The only restriction the streaming API is that not every tweet is returned if the number of matching tweets exceeds 1% of the total volume of tweets.

[6] Download time is a step function. The amount of data returned in each request is small, so Twitter can return the results from 15 requests in a matter of seconds. A user with 75,000 followers therefore requires a few seconds to download the list of followers, but one with 75,001 requires just over 15 minutes.

Third, when asking for an account's previous tweets, the REST API returns only the 3,200 most recent tweets from an account. This restriction means that one will not obtain, for free, all the tweets from accounts which tweet often, such as celebrities, politicians, or media accounts. Because there is a positive relationship between tweet frequency and number of followers [Gonzalez-Bailon, Borge-Holthoefer & Moreno, 2013], the limit means that a sample of users for whom all the tweets are available is a sample of less popular accounts. Moreover, the metadata of each tweet is not reflective of the tweet at the time it was authored. For example, a tweet from the streaming API will show how many followers the author has when the tweet is created, but that same tweet downloaded from the REST API will show the number of followers the account has at the time the old tweet was downloaded. This wrinkle means that tweets from the REST API are not equivalent to tweets from the stream.

Fourth, while the REST API allows one to search historic tweets, the results are only from the previous 7-9 days and not exhaustive of those days. Twitter does not explain how it decides which tweets to return, so it should not be relied upon to reconstruct histories. Twitter only returns 100 results per request, up to 180 requests per 15 minutes.[7] A broad search with thousands of results may therefore take awhile to download and will not provide the population of tweets matching a search query.[8]

Fifth, Twitter only allows one to share 50,000 public tweets and/or accounts' metadata per day, and the sharing cannot be automated. For example, if a researcher uses more than 50,000 tweets for a paper and needs to share them, he or she cannot

[7] Technically, the number of requests depends on whether one is authenticated as a user or application. Since whether or not one form of authentication returns more results depends on the type of request and most academics are not trying to build an application, the rate limits presented are for the user authentication.

[8] Searching directly at www.twitter.com returns all historic matches, but one cannot download those matches.

make them freely available. A system would need to be constructed to verify that an interested party is not downloading the data more than once per day, and the data cannot be pushed to an interested party. Twitter does allow, however, the unlimited distribution of the numeric identification number of each tweet or user account. An interested party can then take these numbers to the REST API and download the full tweet or account.

Sixth, the streaming API occasionally disconnects. These disconnections are rare and random but can imperil one's research design if not caught quickly. Two solutions are available. If one's connection is designed to last indefinitely, that connection's code should generate an e-mail whenever the connection is interrupted. Alternatively, one can intentionally disconnect from the streaming API and immediately reconnect at present intervals, such as every 1 to 24 hours. Any stoppage of the stream would therefore impact only that interval's data collection.

Seventh, when requesting specific users, best practice calls for using an ID number. When a user changes his or her username, Twitter does not update the user ID corresponding to the original screen name. Asking either API for user information based on the screen name may therefore suffer from decay as users change their name, while asking by user ID will not.

4.4.5 A Note on Programming

Twitter data can be acquired, processed, and analyzed with many programming languages, including R and Python. This section briefly describes each language and how they can be used to work with Twitter data. It should be noted, however, that neither is strictly superior to the other, and which to use is as much a matter of personal preferences as anything else.

R is the programming language most familiar to political scientists. An extension of S, a language developed at Bell Laboratories in the 1970s, R was designed by statisticians. While it can perform many general computation functions, its comparative advantage is in statistical analysis. When a new statistical procedure is developed, the first implementation is usually as an R package.

R can ingest Twitter data in 3 ways. streamR is a package that makes it easy to connect to R's streaming API and write the returned tweets to a .csv or JSON file [Barbera, 2013].[9] twitteR is designed to work with the REST API, though the complexity of the REST API compared to the streaming one means the package is not as robust as streamR [Gentry, 2015]. One could avoid these packages completely by using the RCurl package, which facilitates interaction with the HTTP endpoints that web services, including Twitter, use [Lang, 2015]. Using RCurl provides the most flexibility but requires more coding than using a package designed to work with Twitter.

Python is a general purpose language tracing its lineage to 1989 and is most famous for having easy to read code.[10] Whereas R was created for data analysis and has been extended to other purposes, Python was created to work with computers and has been extended to data analysis. Transitioning from writing in R to writing in Python for the first time is much easier than transitioning from never having written code to writing in R.

The primary Python library for working with Twitter is twython [McGrath,

[9] JSON stands for JavaScript Object Notation and is a common formatting protocol for transferring data.

[10] See https://xkcd.com/353/ for a humorous explanation. While any code is harder to read than prose, Python code is as clear, if not more, as R.

2015].[11] Unlike any R package, twython can work with Twitter's REST or streaming API and has built in exception handling. Python's pandas library provides data frames equivalent to R's as well as reshaping, merging, and aggregation capabilities spread across multiple R packages [McKinney, 2015]; pandas is much faster than base R, though R's data.table package is as fast to slightly faster than pandas [Dowle et al., 2015]. Python's statistical libraries are not as deep as R, though most parametric and non-parametric models are available through the statsmodels package [Seabold & Perktold, 2014]. Libraries for Bayesian analysis are not as developed, though Stan has a Python interface. Python has extensive libraries for natural language processing [Bird, 2015] and machine learning [Mueller, 2015] as well.

Neither R nor Python strictly dominates the other. Python has more developed tools for scraping web pages, but Hadley Wickham's rvest package narrows this gap. Python is generally faster, but new R packages such as data.table erase that difference on some dimensions. R's syntax does not resemble other computer languages', but it is also easy to read and learn. The one area Python dominates R for data analysis is data storage: many more database products have Python libraries than R ones, though R has libraries for working with SQL, SQLLite, and MongoDB (a prominent NoSQL database).[12] R dominates Python in developing aesthetically pleasing graphics, though a Python port of ggplot is being developed and Python's matplotlib library produces Matlib style graphics. If one already knows R, learning Python may be worthwhile, but the costs and benefits require careful consideration since either language can most likely accomplish one's programming task and human time is the most scarce resource on any project. The more likely one is to work

[11] tweepy is another Python library to access Twitter, but twython has a larger community and is more frequently updated.

[12] Moreover, most analyzes one will undertake most likely do not require a database.

with large amounts of data or colleagues from outside the social sciences, the more beneficial Python knowledge becomes. If one knows neither R nor Python, learn both.

4.5 Types of Questions

Twitter data can be used to answer questions that involve three kinds of data: networks, text, and spatial. This section considers each in turn.

4.5.1 Networks

Perhaps the most exciting potential of Twitter is as a tool for reconstructing social networks. Networks can be reconstructed from streaming data or data downloaded through the REST API. Because the REST API provides follower and friend information, only it permits the reconstruction of all of a user's connections. A researcher interested in the complete social network of an individual therefore has to use the REST API.

The main downside of the REST API is, again, rate limits. For example, one can only download the follower or friend information for a maximum of 15 people every 15 minutes. Since each request returns up to 5,000 followers or friends, a user with more than 5,000 of either will require multiple requests, and each request counts towards the 15 per 15 minute limit. Since the number of friends or followers grows exponentially as one moves away from a starting user and a researcher is most likely interested in networks of hundreds or thousands of people, the rate limits quickly become become a barrier.

There are 4 tricks one can implement to decrease the crawling time of a network. First, the task can be split across multiple computers. Because downloading the information on each user does not depend on who has been downloaded - there are no interconnections within the follower or friend lists - the task can be assigned to as many computers as possible. Splitting the task across multiple computers is the easiest way to quicken the download process. Second, one can prune any individual with too many followers. This will bias the resulting network away from the core, which may be fine depending on the research question. Third, the researcher can accept only the first 5000 followers or friends of an account. Twitter returns friends and followers in ordered lists of 5,000, where newer connections are returned before older ones. Since each returned list counts towards the rate limit, an account with 75,000 followers therefore takes as long to download as 15 accounts with 5,000 each. In practice, the distribution of followers and friends follows a power law, so most downloading time is concentrated on a subset of accounts with many followers. Restricting a download to only one page of followers will therefore provide the greatest speed increase. Fourth, one can download followers or friends for accounts based on a filtering criterion. For example, if you only care about English speaking accounts, you can remove any accounts whose primary language is different from the list of accounts' for whom you will download followers or friends. This process assumes you already have identifying information on account, which means it is either a seed account or a follower/friend for which you have already downloaded information or you have also downloaded the user information for a list of followers or friends. In other words, you cannot use it to evaluate followers or friends to download, only the followers or friends of those already downloaded.

If one is interested in the strongest connections an account maintains, the streaming data are best. Using the streaming data, the researcher can create edges based on retweets or user mentions. For example, "RT @ZacharyST: Hello world" implies some connection between @ZacharyST and the account retweeting. Similarly, in "@ZacharyST: LeBron James is over-rated", a connection can be inferred between @ZacharyST and the tweet author. Using retweets as an edge will oversample popular users, as they get retweeted the most [Boyd, Golder & Lotan, 2010, Kwak et al., 2010, Suh et al., 2010], while user mentions will produce the fewest edges [Conover et al., 2011, Zamal, Liu & Ruths, 2012*b*]. Using either also assumes that the edge represents an actual following or follower relationship; since a user can retweet any tweet or mention any account, a network based on them may infer links that do not actually exist on Twitter, much less offline.

4.5.2 Text

Office seekers release policy papers. Lawmakers propose and amend bills. Citizens and advocacy groups draft and circulate petitions. Newspapers and television stations make their stories available online, and tools like LexisNexis provide digital access to old stories. Individuals interact verbally, and these interactions are increasingly recorded in text form. Floor speeches and Congressional hearings are quickly released online. Many television shows release digital transcripts. Computationally analyzing these data has quickly become a growing frontier in political science [Grimmer & Stewart, 2013, Lucas et al., 2015].

The barriers to entry for Twitter are substantially lower than for appearing in a news report or in Congress. People who are not likely to appear in those sources are therefore much more likely to appear on Twitter, casting light on vast swathes of previously unobservable behavior. While this outcome is true of the internet more broadly - barriers to entry are similarly low for joining most social networks, posting to web forums, or starting a blog - Twitter is unique in its combination of size and public communication. Elites who are likely to appear in traditional text sources usually have Twitter accounts (every Senator and 430 members of the House have official Twitter accounts, as do most news organization), meaning Twitter records communication behavior from all strata of society. No other data source exists which is simultaneously comprehensive and accessible.

Studies have used text to examine a variety of phenomena. Tweets can be used to measure sentiment across time and space in ways consistent with time of day and season [Dodds et al., 2011, Golder & Macy, 2011] as well as distance from home [Frank et al., 2013], and this sentiment can predict changes in the Dow Jones Industrial Average [Bollen, Mao & Zeng, 2011] and the Standard & Poor 500 [Zheludev, Smith & Aste, 2014]. Textual analysis also suggests that members of Congress' rhetoric follows that of those they represent, not the other way around [Barberá et al., 2014]. Tweets can also reveal dynamic patterns of political polarization in countries as diverse as the United States [Barberá, Jost, Nagler, Tucker & Bonneau, 2015] and Egypt [Borge-Holthoefer et al., 2015]. Text in tweets can also be used to predict upcoming crimes [Gerber, 2014] and is used as one of many textual inputs to predict political unrest [Kallus, 2013]. Studies of Twitter during local emergencies suggest that information is shared in a two-step flow model [Katz, 1957], with most people sharing information they first obtained from news organizations [Starbird & Palen, 2010, Vieweg et al., 2010, Yardi & Boyd, 2010].

There are two idiosyncrasies to working with tweet text. First, Twitter requires all messages to contain less than 140 characters. The benefit of this limitation is that it tweets tend to concern themselves with one topic, simplifying analysis. The drawback of this limitation is that it pushes users to use abbreviations and slang, which complicate language processing. Second, tweet style is bimodal, with very many, perhaps a majority, of them using abbreviations and slang. Existing corpora used for dictionary approaches do not include slang, and the idiosyncratic nature of slang means unsupervised approaches are more likely to assign tweets about the same topic to different collections. For a more detailed discussion on natural language processing and Twitter, see Sriram et. al. (2010) and Han and Baldwin (2011).[13]

4.5.3 Spatial

It is also possible to map Twitter activity to specific places, allowing scholars to connect patterns on Twitter to offline events. Location information comes from two sources, accounts choosing to provide their GPS coordinates or self-reporting their location as part of their profile.

Tweets with GPS coordinates, which represent 2-3% of all tweets [Leetaru et al., 2013], provide the most precision in estimating location. They can be used to map language communities at the city and neighborhood level [Hale, Gaffney & Graham, 2011, Mocanu et al., 2013] and unemployment [Llorente et al., 2014]. They also reveal that offline geography conditions online network connections [Yardi & Boyd, 2010, Kulshrestha et al., 2012, Conover et al., 2013]. In addition to representing few tweets, they are more prevalent in urban areas and among higher-income users,

[13] For a thorough introduction to natural language processing more broadly, see Manning and Schütze (1999).

so the extent to which they are reliable depends on the research question [Malik et al., 2015].

To increase the number of geolocated tweets in a sample, scholars infer location based on a user's self-reported location in their profile, though the user does not have to identify their location. This location can be any text string, so some users report their location as "Space" or "Under the Sea", but most reported locations correspond to actual places. Including users' self-reported location, anywhere from 34% [Leetaru et al., 2013] to 50% [Conover et al., 2013] to 66% [Hecht et al., 2011] of tweets have location information, the vast majority of which are at the city level. It is possible to assign location information to accounts based on the context of their tweets [Cheng, Caverlee & Lee, 2010, Hecht et al., 2011, Stefanidis, Crooks & Radzikowski, 2011]. It is unknown whether accounts with self-reported location differ from those without, so the benefit of this added parsing needs to weighed against the computational cost. The most global source of location is an account's time zone, as Twitter automatically assigns a time zone to each tweet based on the location where the tweet author registered [Lotan et al., 2011].

4.6 Conclusion

The rise of digital communications data has pushed the boundaries of scholarship. Traditionally, whether a researcher pursues qualitative, case-study focused methods or quantitative, regression-based techniques depends on the research question. If the researcher is interested in a particular part of the world or the evidence required to test a theory is too complex to be quantified, qualitative approaches are favored. If the researcher is more concerned about patterns across a wide range of cases and theory predicts the main effects will not be lost in data aggregation, then quantitative approaches are followed.

The rise of online data lowers the walls between these two approaches. For example, studies of protest and repression have either focused on one or a couple of countries in order to quantify temporal patterns at the daily or weekly level [Lohmann, 1994, Rasler, 1996, Francisco, 1996] or aggregate across many countries at a monthly or yearly frequency to try and reach more general conclusions [Gurr, 1993, Danneman & Ritter, 2013]. The advent of near real-time events datasets and social networks, however, has allowed researchers to study protest dynamics at a daily level [González-Bailón et al., 2011, Lotan et al., 2011, King, Pan & Roberts, 2013] and across multiple countries. This new generation of work has allowed scholars of a quantitative orientation to gather more data about particular cases and makes it easier for qualitative-focused ones to compare their cases to a wider range of cases. Both approaches' theories improve as a result of these data.

These data are not a "revolution". Instead, they represent the next stage in the constant increase in data available to researchers. In the 1970s, cutting-edge empirical research consisted of descriptive statistics and basic regressions; using a computer required access to a mainframe and programming with punch cards. The 1980s and 1990s democratized computers for professionals, though data analysis programs did not become usable to those without programming knowledge until the 1990s. The last two decades have witnessed the rise of digital communications data. Because these data are much cheaper to produce than human-created datasets, the amount of data available for analysis has expanded. This expansion has made previous approaches - aggregating data to the national or annual level, modeling data through graphical user interfaces such as Stata or SPSS, loading an entire dataset at once, &c - often inadequate. Now, to stay at the forefront of data analysis, one needs to know some programming in order to interface with websites and data services, download data automatically, algorithmically clean and analyze data, and present these data in low-dimension environments. The skills are modern; the change is eternal.

4.7 Acknowledgements

Chapter 4, in part, has been submitted for publication. I am the primary investigator and author of this paper.

Bibliography

Abramson, Paul R. & William Clagget. 2001. "Recruitment and Political Participation." *Political Research Quarterly* 54(4):905–916.

Aday, Sean, Deen Freelon, Henry Farrell, Marc Lynch & John Sides. 2012. New Media and Conflict After the Arab Spring. Technical report United States Institute of Peace Washington D.C.: .

Al-Zubaidi, Layla & Matthew Cassel, eds. 2013. *Diaries of an Unfinished Revolution: Voices from Tunis to Damascus.* Penguin Books.

Albrecht, Glenn. 2005. "Solastalgia: A New Concept in Health and Identity." (3):41–55.

Amar, Paul. 2011. "Turning the Gendered Politics of the Security State Inside Out? Charging the Police with Sexual Harassment in Egypt." *International Feminist Journal of Politics* 13(3):299–328.

Andrews, K. T. & M. Biggs. 2006. "The Dynamics of Protest Diffusion: Movement Organizations, Social Networks, and News Media in the 1960 Sit-Ins." *American Sociological Review* 71(5):752–777.

Aral, Sinan & Dylan Walker. 2012. "Identifying Influential and Susceptible Members of Social Networks." *Science* 337(July):337–341.

Arceneaux, K. & D. W. Nickerson. 2009. "Modeling Certainty with Clustered Data: A Comparison of Methods." *Political Analysis* 17(2):177–190.
URL: *http://pan.oxfordjournals.org/cgi/doi/10.1093/pan/mpp004*

Arva, Bryan, John Beieler, Ben Fisher, Gustavo Lara, Philip A Schrodt, Wonjun Song, Marsha Sowell & Sam Stehle. 2013. "Improving Forecasts of International Events of Interest.".

Asur, Sitaram & Bernardo A. Huberman. 2010. Predicting the Future with Social Media. In *2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology.* Ieee pp. 492–499.

Bail, Christopher A. 2014. "The cultural environment: measuring culture with big data." *Theory and Society* 43(3-4):465–482.

Barbera, Pablo. 2013. "streamR.".
URL: *https://cran.r-project.org/web/packages/streamR/*

Barbera, Pablo. 2015. "Birds of the Same Feather Tweet Together: Bayesian Ideal Point Estimation Using Twitter Data." *Political Analysis* 23(August 2013):76–91.

Barberá, Pablo, John T Jost, Jonathan Nagler, Joshua a Tucker & Richard Bonneau. 2015. "Tweeting From Left to Right: Is Online Political Communication More Than an Echo Chamber?" *Psychological science* pp. 1–12.

Barberá, Pablo, Ning Wang, Richard Bonneau, John T. Jost, Jonathan Nagler, Joshua Tucker & Sandra González-Bailón. 2015. "The Critical Periphery in the Growth of Social Protests." *PloS ONE* pp. 1–15.

Barberá, Pablo, Richard Bonneau, Patrick Egan, John T. Jost, Jonathan Nagler & Joshua Tucker. 2014. "Leaders or Followers? Measuring Political Responsiveness in the U.S. Congress Using Social Media Data.".

Bassiouni, Mahmoud Cherif, Nigel Rodley, Badria Al-Awadhi, Philippe Kirsch & Mahnoush H. Arsanjani. 2011. Report of the Bahrain Independent Commission of Inquiry. Technical Report NOVEMBER Bahrain Independent Commission of Inquiry Manama, Bahrain: .

Beinin, Joel. 2009. Workers' Struggles Under 'Socialism' and Neoliberalism. In *Egypt: The Moment of Change*, ed. Philip Marfleet & Rabab El-Mahdi. London: Zed Books pp. 68–86.

Beissinger, Mark R. 2013. "The Semblance of Democratic Revolution: Coalitions in Ukraine's Orange Revolution." *American Political Science Review* 107(03):574–592.

Beissinger, Mark R., Amaney A. Jamal & Kevin Mazur. 2015. "Explaining Divergent Revolutionary Coalitions: Regime Strategies and the Structuring of Participation in the Tunisian and Egyptian Revolutions." *Comparative Politics* 48(1):1–24.

Bellin, Eva. 2012. "Reconsidering the Robustness of Authoritarianism in the Middle East: Lessons from the Arab Spring." *Comparative Political Studies* 44(2):127–149.

Benchemsi, Ahmed. 2014. Morocco's Makhzen and the Haphazard Activists. In *Taking to the Streets: The Transformationof Arab Activism*, ed. Lina Khatib & Ellen Lust. Baltimore: Johns Hopkins University Press pp. 199–236.

Bennett, W. Lance & Alexandra Segerberg. 2013. *The Logic of Connective Action*. Number June 2013 Cambridge: Cambridge University Press.

Bird, Steven. 2015. "Natural Language Toolkit.".

Bishara, Dina. 2014. The Power of Workers in Egypt's 2011 Uprising. In *Arab Spring in Egypt: Revolution and Beyond*, ed. Bahgat Korany & Rabab El-Mahdi. AUC Press pp. 83–105.

Blaydes, L. & J. Lo. 2011. "One man, one vote, one time? A model of democratization in the Middle East." *Journal of Theoretical Politics* 24(1):110–146.

Bollen, Johan, Huina Mao & Xiaojun Zeng. 2011. "Twitter mood predicts the stock market." *Journal of Computational Science* 2(1):1–8.

Bond, Robert M., Christopher J. Fariss, Jason J. Jones, Adam D.I. Kramer, Cameron Marlow, Jaime E. Settle & James H Fowler. 2012. "A 61-million-person experiment in social influence and political mobilization." *Nature* 489(7415):295–8.

Borge-Holthoefer, Javier, Walid Magdy, Kareem Darwish & Ingmar Weber. 2015. Content and Network Dynamics Behind Egyptian Political Polarization on Twitter. In *18th Conference on Computer-Supported Cooperative Work and Social Computing*. pp. 1–30.

Boschee, Elizabeth, Jennifer Lautenschlager, Sean O'Brien, Steve Shellman, James Starz & Michael Ward. 2015. "ICEWS Coded Event Data.".
URL: *http://dx.doi.org/10.7910/DVN/28075*

Boudreau, Vincent. 2009. *Resisting Dictatorship: Repression and Protest in Southeast Asia*. Cambridge: Cambridge University Press.

Boyd, Dannah, Scott Golder & Gilad Lotan. 2010. Tweet, Tweet, Retweet: Conversational Aspects of Retweeting on Twitter. In *43rd Hawaii International Conference on System Sciences*. Ieee pp. 1–10.

Brady, Henry E., Kay Lehman Schlozman & Sidney Verba. 1999. "Prospecting for Participants: Rational Expectations and the Recruitment of Political Activists." *The American Political Science Review* 93(1):153–168.

Breuer, Anita. 2012. The Role of social media in mobilizing political protest: Evidence from the Tunisian Revolution. Technical report German Development Institute Bonn: .

Broder, Andrei, Ravi Kumar, Farzin Maghoul, Prabhakar Raghavan, Sridhar Rajagopalan, Raymie Stata, Andrew Tomkins & Janet Wiener. 2000. "Graph structure in the Web." *Computer Networks* 33(1-6):309–320.

Bruns, Axel & Jean E Burgess. 2011. The Use of Twitter Hashtags in the Formation of Ad Hoc Publics. In *6th European Consortium for Political Research Conferenc*. Number August Reykjavik: pp. 1–10.

Bueno de Mesquita, Bruce, Alastair Smith, Randolph M. Siverson & James D. Morrow. 2003. *The Logic of Political Survival*. Cambridge: MIT Press.

Burns, Alex & Ben Eltham. 2009. Twitter Free Iran: An Evaluation of Twitter's Role in Public Diplomacy and Information Operations in Iran's 2009 Election Crisis. In *Communications Policy & Research Forum 2009*. Number February 2008 Sydney: pp. 298–310.

Cambanis, Thanassis. 2015. *Once Upon a Revolution: An Egyptian Story*. New York: Simon and Schuster.

Cameron, Ann Frances & Jane Webster. 2005. "Unintended consequences of emerging communication technologies: Instant Messaging in the workplace." *Computers in Human Behavior* 21(1):85–103.

Caren, Neal. 2014. "It is time to get rid of the E in GDELT.".
URL: *http://badhessian.org/2014/05/it-is-time-to-get-rid-of-the-e-in-gdelt/*

Carey, Sabine C. 2010. "The Use of Repression as a Response to Domestic Dissent." *Political Studies* 58(1):167–186.

Carlsson, Hans & Eric van Damme. 1993. "Global Games and Equilibrium Selection.".

Carpenter, R. Charli. 2011. "Vetting the Advocacy Agenda: Network Centrality and the Paradox of Weapons Norms." *International Organization* 65(01):69–102.

Centola, Damon. 2010. "The Spread of Behavior in an Online Social Network Experiment." *Science* 329(September):1194–1198.

Centola, Damon & Michael Macy. 2007. "Complex Contagion and the Weakness of Long Ties." *American Journal of Sociology* 113(3):702–734.

Cheng, Zhiyuan, James Caverlee & Kyumin Lee. 2010. You Are Where You Tweet: A Content-Based Approach to Geo-locating Twitter Users. In *ACM International Conference on Information and Knowledge Management*. Toronto: .

Chenoweth, E. & J. Ulfelder. 2015. "Can Structural Conditions Explain the Onset of Nonviolent Uprisings?" *Journal of Conflict Resolution* pp. 1–27.

Chew, Cynthia & Gunther Eysenbach. 2010. "Pandemics in the age of Twitter: Content Analysis of Tweets during the 2009 H1N1 Outbreak." *PloS one* 5(11):e14118.

Chomiak, Laryssa. 2014. Architecture of Resistance in Tunisia. In *Taking to the Streets: The Transformationof Arab Activism*, ed. Lina Khatib & Ellen Lust. Baltimore: Johns Hopkins University Press chapter 1, pp. 22–52.

Christakis, Nicholas A. & James H. Fowler. 2012. "Social contagion theory: examining dynamic social networks and human behavior." *Statistics in Medicine* 32(4):556–577.

Chwe, Michael. 1998. "Believe the Hype: Solving Coordination Problems with Television Advertising.".

Chwe, Michael Suk-young. 2003. *Rational Ritual: Culture, Coordination, and Common Knowledge*. Princeton: Princeton University Press.

Citino, Robert. 2004. "Beyond Fire and Movement: Command, Control and Information in the German Blitzkrieg." *Journal of Strategic Studies* 27(2):324–344.

Conover, M.D., J. Ratkiewicz, M. Francisco, B. Goncalves, A. Flammini & F. Menczer. 2011. Political Polarization on Twitter. In *Fifth International AAAI Conference on Weblogs and Social Media*. pp. 89–96.

Conover, Michael D, Clayton Davis, Emilio Ferrara, Karissa McKelvey, Filippo Menczer & Alessandro Flammini. 2013. "The geospatial characteristics of a social movement communication network." *PloS one* 8(3):e55957.

Dalton, Russell J., Steven Greene, Paul Allen Beck & Robert Huckfeldt. 2002. "The Social Calculus of Voting: Interpersonal, Media, and Organizational Influences on Presidential Choices." *The American Political Science Review* 96(1):57–73.

Danneman, Nathan & Emily H. Ritter. 2013. "Contagious Rebellion and Preemptive Repression." *Journal of Conflict Resolution* 58(2):254–279.

Davenport, Christian. 2007*a*. "State Repression and Political Order." *Annual Review of Political Science* 10(1):1–23.

Davenport, Christian. 2007*b*. "State Repression and Political Order." *Annual Review of Political Science* 10(1):1–23.

Davenport, Christian. 2007*c*. *State Repression and the Domestic Democratic Peace*. Cambridge: Cambridge University Press.

Davenport, Christian & Patrick Ball. 2002. "Views to a Kill: Exploring the Implications of Source Selection in the Case of Guatemalan State Terror, 1977-1995." *Journal of Conflict Resolution* 46(3):427–450.

Davenport, Christian & Priyamvada Trivedi. 2013. "Activism and awareness: Resistance, cognitive activation, and 'seeing' untouchability among 98,316 Dalits." *Journal of Peace Research* 50(3):369–383.

De Mesquita, Ethan Bueno. 2010. "Regime Change and Revolutionary Entrepreneurs." *American Political Science Review* 104(03):446–466.

Dodds, Peter Sheridan, Kameron Decker Harris, Isabel M. Kloumann, Catherine A. Bliss & Christopher M. Danforth. 2011. "Temporal Patterns of Happiness and Information in a Global Social Network: Hedonometrics and Twitter." *PLoS ONEcomput* 6(12):e26752.

Donner, Allan. 1998. "Some aspects of the design and analysis of cluster randomization trials." *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 47(1):95–113.

Dowle, Matt, T Short, S Lianoglou & A Srinivasan. 2015. "data.table: Extension of data.frame.".
URL: *https://cran.r-project.org/web/packages/data.table/index.html*

Downs, Anthony. 1957. *An Economic Theory of Democracy.* New York City: Harper and Row.

Duggan, Maeve & Aaron Smith. 2013. Social Media Update 2013. Technical Report January Pew Research Center.

Dunbar, R.I.M., Valerio Arnaboldi, Marco Conti & Andrea Passarella. 2015. "The structure of online social networks mirrors those in the offline world." *Social Networks* 43:39–47.

Earl, Jennifer, Heather McKee Hurwitz, Analicia Mejia Mesinas, Margaret Tolan & Ashley Arlotti. 2013. "This Protest Will Be Tweeted: Twitter and protest policing during the Pittsburgh G20." *Information, Communication & Society* 16(4):459–478.

Ebaid, Neama. 2013. Sexual Harassment in Egypt: A Neglected Crime. Technical report Academy for Cultural Diplomacy Cairo: .

Eck, Kristine. 2012. "In data we trust? A comparison of UCDP GED and ACLED conflict events datasets." *Cooperation and Conflict* 47(1):124–141.

Edmond, Chris. 2013. "Information Manipulation, Coordination, and Regime Change." *Review of Economic Studies* 80(4):1422–1458.

Egorov, Georgy, Sergei Guriev & Konstantin Sonin. 2009. "Why Resource-poor Dictators Allow Freer Media: A Theory and Evidence from Panel Data." *American Political Science Review* 103(04):645.

Fearon, J. D. 2011. "Self-Enforcing Democracy." *The Quarterly Journal of Economics* 126(4):1661–1708.

Fearon, James D. & David D. Laitin. 2003. "Ethnicity, Insurgency, and Civil War." *The American Political Science Review* 97(1):75–90.

Ferrara, F. 2003. "Why Regimes Create Disorder: Hobbes's Dilemma During a Rangoon Summer." *Journal of Conflict Resolution* 47(3):302–325.

Finkel, Steven E., Edward N. Muller & Karl-Dieter Opp. 1989. "Personal Influence, Collective Rationality, and Mass Political Action." *American Political Science Review* 83(3):885–903.

Francisco, Ronald A. 1996. "Coercion and Protest: An Empirical Test in Two Democratic States." *Journal of Political Science* 40(4):1179–1204.

Francisco, Ronald A. 2004. "After the Massacre: Mobilization in the Wake of Harsh Repression." *Mobilization: An International Journal* 9(2):107–126.

Frank, Morgan R, Lewis Mitchell, Peter Sheridan Dodds & Christopher M Danforth. 2013. "Happiness and the patterns of life: a study of geolocated tweets." *Scientific reports* 3:2625.

Freelon, Dean. 2012. "Arab Spring Twitter data now available (sort of).".
URL: *http://dfreelon.org/2012/02/11/arab-spring-twitter-data-now-available-sort-of/*

Gabielkov, Maksym, Ashwin Rao & Arnauld Legout. 2014. "Studying Social Networks at Scale : Macroscopic Anatomy of the Twitter Social Graph.".

Gandhi, Jennifer & Ellen Lust-Okar. 2009. "Elections Under Authoritarianism." *Annual Review of Political Science* 12(1):403–422.

Garcia-Herranz, Manuel, Esteban Moro, Manuel Cebrian, Nicholas A. Christakis & James H. Fowler. 2014. "Using Friends as Sensors to Detect Global-Scale Contagious Outbreaks." *PloS ONE* 9(4):e92413.

Gayo-Avello, Daniel. 2013. "A Meta-Analysis of State-of-the-Art Electoral Prediction From Twitter Data." *Social Science Computer Review* 31(6):649–679.

Gentry, Jeff. 2015. "twitteR.".
URL: *https://cran.r-project.org/web/packages/twitteR/index.html*

Gerbaudo, Paolo. 2012. *Tweets and the Streets: Social Media and Contemporary Activism*. London: Pluto Press.

Gerber, Alan, Dean Karlan & Daniel Bergan. 2006. "Does The Media Matter? A Field Experiment Measuring the Effect of Newspapers on Voting Behavior and Political Opinions.".

Gerber, Alan S., Donald P. Green & Christopher W. Larimer. 2008. "Social Pressure and Voter Turnout: Evidence from a Large-Scale Field Experiment." *American Political Science Review* 102(01):33–48.

Gerber, Matthew S. 2014. "Predicting crime using Twitter and kernel density estimation." *Decision Support Systems* 61:115–125.

Gerner, Deborah J., Philip A. Schrodt, Rajaa Abu-Jabr & Omur Yilmaz. 2002. Conflict and Mediation Event Observations (CAMEO): A New Event Data Framework for the Analysis of Foreign Policy Interactions. In *Annual Meeting of the International Studies Association*.

Ghonim, Wael. 2012. *Revolution 2.0: The Power of the People is Greater Than the People in Power: A Memoir*. Houghton Mifflin Co.

Gilbert, Eric. 2013. Widespread Underprovision on Reddit. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*. New York, New York, USA: ACM Press p. 803.
URL: *http://dl.acm.org/citation.cfm?doid=2441776.2441866*

Glanz, James & John Markoff. 2011. "Egypt Leaders Found 'Off' Switch for Internet.".
URL: *http://www.nytimes.com/2011/02/16/technology/16internet.html?_r=1&hp=&pagewanted=all*

Golder, Scott A & Michael W Macy. 2011. "Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures." *Science (New York, N.Y.)* 333(6051):1878–81.

Goldstone, Jack A. 1994. "Is Revolution Individually Rational?: Groups and Individuals in Revolutionary Collective Action." *Rationality and Society* 6(1):139–166.

Goldstone, Jack A. 2001. "Toward a Fourth Generation of Revolutionary Theory." *Annual Review of Political Science* 4:139–187.

González-Bailón, Sandra, Javier Borge-Holthoefer, Alejandro Rivero & Yamir Moreno. 2011. "The Dynamics of Protest Recruitment through an Online Network." *Scientific Reports* 1:197.

Gonzalez-Bailon, Sandra, Javier Borge-Holthoefer & Yamir Moreno. 2013. "Broadcasters and Hidden Influentials in Online Protest Diffusion." *American Behavioral Scientist* 57(7):943–965.

González-Bailón, Sandra, Ning Wang, Alejandro Rivero, Javier Borge-Holthoefer & Yamir Moreno. 2012. "Assessing the Bias in Communication Networks Sampled from Twitter.".

Granovetter, Mark. 1973. "The Strength of Weak Ties." *American Journal of Sociology* 78:1360–1380.

Granovetter, Mark. 1978. "Threshold Models of Collective Behavior." *American Journal of Sociology* 83(6):1420–1443.

Grimmer, J. & B. M. Stewart. 2013. "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts." *Political Analysis* 21(3):267–297.

Grimmer, Justin. 2015. "We Are All Social Scientists Now: How Big Data, Machine Learning, and Causal Inference Work Together." *PS: Political Science & Politics* 48(01):80–83.

Groshek, Jacob. 2015. "Status Update on the BU-TCAT.".
URL: *http://www.jgroshek.org/blog/2015/8/17/status-update-on-the-bu-tcat*

Gunitsky, Seva. 2015. "Corrupting the Cyber-Commons: Social Media as a Tool of Autocratic Stability." *Perspectives on Politics* 13(01):42–54.

Gunning, Jeroen & Ilan Zvi Baron. 2013. *Why Occupy a Square: People, Protests and Movements in the Egyptian Revolution*. London: Hurst Publishers.

Gurr, T. R. 1993. "Why Minorities Rebel: A Global Analysis of Communal Mobilization and Conflict since 1945." *International Political Science Review* 14(2):161–201.

Gurr, Ted. 1971. *Why Men Rebel*. Princeton: Princeton University Press.

Hale, Henry E. 2013. "Regime Change Cascades: What We Have Learned from the 1848 Revolutions to the 2011 Arab Uprisings." *Annual Review of Political Science* 16(1):331–353.

Hale, Scott A., Devin Gaffney & Mark Graham. 2011. "Where in the world are you? Geolocation and language identification in Twitter.".

Hamdy, Naila & Ehab H. Gomaa. 2012. "Framing the Egyptian Uprising in Arabic Language Newspapers and Social Media." *Journal of Communication* 62(2):195–211.

Hammond, Jesse & Nils B. Weidmann. 2014. "Using machine-coded event data for the micro-level study of political violence." *Research & Politics* 1(2):1–8.

Hanna, Alexander. 2013. "Computer-aided Content Analysis of Digitally Enabled Movements." *Mobilization: An International Quarterly* 18(4):367–388.

Hassanpour, Navid. 2014. "Media Disruption and Revolutionary Unrest: Evidence From Mubarak's Quasi-Experiment." *Political Communication* 31(1):1–24.

Hecht, Brent, Lichan Hong, Bongwon Suh & Ed H. Chi. 2011. Tweets from Justin Biebers Heart: The Dynamics of the Location Field in User Profiles. In *ACM Conference on Human Factors in Computing Systems*. Number Figure 1 Vancouver: .

Hendrix, Cullen, Jesse Hamner, Christina Case, Chris Linebarger, Emily Stull & Jennifer Williams. 2012. "Social Conflict in Africa: A New Database." *International Interactions* 38(4):503–511.

Herkenrath, M. & A. Knoll. 2011. "Protest events in international press coverage: An empirical critique of cross-national conflict databases." *International Journal of Comparative Sociology* 52(3):163–180.

Herring, Susan C., Elijah Wright, Lois Ann Schedit & Sabrina Bonus. 2004. Bridging the Gap: A Genre Analysis of Weblogs. In *37th Hawaii International Conference on System Sciences*. Vol. 00 pp. 1–11.

Hill Jr., Daniel W. & Zachary M. Jones. 2014. "An Empirical Evaluation of Explanations for State Repression." *American Political Science Review* 108(03):661–687.

Hinduja, Sameer & Justin W Patchin. 2008. "Personal information of adolescents on the Internet: A quantitative content analysis of MySpace." *Journal of Adolescence* 31(1):125–46.

Hoffman, Mike. 2015. "US Air Force Targets and Destroys ISIS HQ Building Using Social Media.".

Hollyer, James R., B. Peter Rosendorff & James Raymond Vreeland. 2015. "Transparency, Protest, and Autocratic Instability." *American Political Science Review* 109(04):764–784.

Howard, Philip N., Aiden Duffy, Deen Freelon, Muzammil Hussain, Will Mari & Marwa Mazaid. 2011. Opening Close Regimes: What Was the Role of Social Media During the Arab Spring? Technical report Project on Information Technology & Political Islam.

Howard, Philip N. & Muzammil M. Hussain. 2011. "The Role of Digital Media." *Journal of Democracy* 22(3):35–48.

Human Rights Watch. 2011. World Report 2011: Bahrain. Technical report Human Rights Watch.

Idle, Nadia & Alex Nunns, eds. 2011. *Tweets from Tahrir*. New York City: OR Books.

International Telecommunications Union Statistics. 2014.
URL: *http://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx*

Javeline, Debra. 2003. "The Role of Blame in Collective Action: Evidence from Russia." *American Political Science Review* 97(01):107.

Jones, Toby C. & Alaa Shehabi. 2012. "Bahrains revolutionaries." *Foreign Policy* pp. 12–16.

Kallus, Nathan. 2013. Predicting Crowd Behavior with Big Public Data. In *23rd International Conference on World Wide Web*.

Katz, Elihu. 1957. "The Two-Step Flow of Communication: An Up-To-Date Report on an Hypothesis." *Public Opinion Quarterly* 21(1):61–78.

Keck, Margaret E. & Kathryn Sikkink. 1998. *Activists Beyond Borders: Advocacy Networks in International Politics*. Ithaca: Cornell University Press.

Kerr, Simeon. 2015. "Bahrain main Shia opposition leader jailed for 4 years.".

Khalil, Ashraf. 2011. *Liberation Square: Inside the Egyptian Revolution and the Rebirth of a Nation*. New York City: St. Martin's Press.

Khatib, Lina & Ellen Lust, eds. 2014. *Taking to the Streets: The Transformation of Arab Activism*. Baltimore: Johns Hopkins University Press.

Kim, Hyojoung & Peter S Bearman. 1997. "The Structure and Dynamics of Movement Participation." *American Sociological Review* 62(1):70–93.

King, Gary, Jennifer Pan & Margaret E. Roberts. 2013. "How Censorship in China Allows Government Criticism but Silences Collective Expression." *American Political Science Review* 107(02):326–343.

Klandermans, Bert & Dirk Oegema. 1987. "Potentials, Networks, Motivations, and Barriers: Steps Towards Participation in Social Movements." *American Sociological Review* 52(4):519–531.

Klofstad, Casey A. 2007. "Talk Leads to Recruitment: How Discussions about Politics and Current Events Increase Civic Participation." *Political Research Quarterly* 60(2):180–191.

Kneeland, Terri. 2014. "Coordination under Limited Depth of Reasoning.".

Kramer, Adam D.I., Jamie E. Guillory & Jeffrey T. Hancock. 2014. "Experimental evidence of massive-scale emotional contagion through social networks." *Proceedings of the National Academy of Sciences* 111(24):8788–8790.

Kulshrestha, Juhi, Farshad Kooti, Ashkan Nikravesh & Krishna P Gummadi. 2012. Geographic Dissection of the Twitter Network. In *Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media*. pp. 202–209.

Kuran, Timur. 1989. "Sparks and Prairie Fires: A Theory of Unanticipated Political Revolution." *Public Choice* 61(1):41–74.

Kuran, Timur. 1991. "Now Out of Never: The Element of Surprise in the East European Revolution of 1989." *World Politics* 44(1):7–48.

Kwak, Haewoon, Changhyun Lee, Hosung Park & Sue Moon. 2010. What is Twitter, a Social Network or a News Media? In *International World Wide Conference*. pp. 591–600.

Lang, Duncam Temple. 2015. "RCurl.".

Langohr, Vickie. 2013. “”This is Our Square”: Fighting Sexual Assault at Cairo Protests.” *Middle East Research and Information Project* 43(Fall).
URL: *http://www.merip.org/mer/mer268/our-square*

Lee, Sang, Pan-Jun Kim & Hawoong Jeong. 2006. “Statistical properties of sampled networks.” *Physical Review E* 73(1):016102.

Leetaru, Kalev H., Shaowen Wang, Guofeng Cao, Anand Padmanabhan & Eric Shook. 2013. “Mapping the global Twitter heartbeat: The geography of Twitter.” *First Monday* 18(5-6):1–33.

Leetaru, Kalev & Philip Schrodt. 2013. “GDELT: Global Data on Events, Language, and Tone, 1979-2012.”.

Lehmann, Janette, Bruno Gonçalves, José J. Ramasco & Ciro Cattuto. 2012. “Dynamical Classes of Collective Attention in Twitter.” *Proceedings of the 21st International Conference on World Wide Web - WWW ’12* pp. 251–260.

Lerman, Kristina & Rumi Ghosh. 2010. Information Contagion: An Empirical Study of Information Spread on Digg and Twitter Follower Graphs. In *Fourth International AAAI Conference on Weblogs and Social Media*.

Leskovec, Jure & Eric Horvitz. 2007. “Worldwide Buzz : Planetary-Scale Views on an Instant-Messaging Network.”.

Li, Dahui, Patrick Y.K. Chau & Hao Lou. 2005. “Understanding Individual Adoption of Instant Messaging :.” *Journal of the Association for Information Systems* 6(4):102–129.

Lichbach, M. I. 1987. “Deterrence or Escalation?: The Puzzle of Aggregate Studies of Repression and Dissent.” *Journal of Conflict Resolution* 31(2):266–297.

Lichbach, Mark I. 1998. “Contending Theories of Contentious Politics and the Structure-Action Problem of Social Order.” *Annual Review of Political Science* 1:401–424.

Liere, Diederik Van. 2010. How Far Does a Tweet Travel? Information Brokers in the Twitterverse. In *Modeling Social Media*. Toronto: pp. 1–4.

Lim, Merlyna. 2012. “Clicks, Cabs, and Coffee Houses: Social Media and Oppositional Movements in Egypt, 2004-2011.” *Journal of Communication* 62(2):231–248.

Llorente, Alejandro, Manuel Garcia-Herranz, Manuel Cebrian & Esteban Moro. 2014. “Social media fingerprints of unemployment.”.
URL: *http://arxiv.org/abs/1411.3140*

Lohmann, Susanne. 1994. “The Dynamics of Informational Cascades: The Monday Demonstrations in Leipzig, East Germany, 1989-91.” *World Politics* 47(1):42–101.

Lotan, Gilad, Mike Ananny, Devin Gaffney, Danah Boyd, Ian Pearce & Erhardt Graeff. 2011. “The Revolutions Were Tweeted: Information Flows During the 2011 Tunisian and Egyptian Revolutions Web.” *International Journal of Communications* 5:1375–1406.

Lucas, Christopher, Richard A. Nielsen, Margaret E. Roberts, Brandon M. Stewart, Alex Storer & Dustin Tingley. 2015. “Computer-Assisted Text Analysis for Comparative Politics.” *Political Analysis* 23(2):254–277.

Mackey, Robert. 2015. “Egyptian Police Attack Marchers Carrying Flowers to Tahrir Square, Killing Protester.”.
URL: *http://www.nytimes.com/2015/01/25/world/middleeast/egypt-woman-is-killed-in-peaceful-cairo-protest.html*

Malik, Momin M, Hemank Lamba, Constantine Nakos & Jurgen Pfeffer. 2015. Population Bias in Geotagged Tweets. In *9th International AAAI Conference on Weblogs and Social Media*. pp. 18–27.

Marwell, Gerald, Pamela E. Oliver & Ralph Prahl. 1988. “Social Networks and Collective Action: A Theory of the Critical Mass.” *American Journal of Sociology* 94(3):502–534.

Masad, David. 2013. “Studying the Syrian Civil War with GDELT.”.

McAdam, Doug. 1986. “Recruitment to High-Risk Activism: The Case of Freedom Summer.” *American Journal of Sociology* 92(1):64–90.

McAdam, Doug & Ronnelle Paulsen. 1993. “Specifying the Relationship between Social Ties and Activism.” *American Journal of Sociology* 99(3):640–667.

McAdam, Doug, Sidney Tarrow & Charles Tilly. 2001. *Dynamics of Contention*. Cambridge: Cambridge University Press.

McGrath, Ryan. 2015. “twython.”.

McKinney, Wes. 2015. “pandas.”.
URL: *http://pandas.pydata.org/*

McPherson, Miller, Lynn Smith-Lovin & James M. Cook. 2001. “Birds of a Feather: Homophily in Social Networks.” *Annual Review of Sociology* 27(2001):415–444.

Meirowitz, Adam & Joshua A. Tucker. 2013. “People Power or a One-Shot Deal? A Dynamic Model of Protest.” *American Journal of Political Science* 00(0):1–13‘.

Miguel, Carolina e, Amaney A. Jamal & Mark Tessler. 2015. “Elections in the Arab World: Why Do Citizens Turn Out ?” *Comparative Political Studies* 48(11):1–34.

Miller, M. K. 2014. "Elections, Information, and Policy Responsiveness in Autocratic Regimes." *Comparative Political Studies* 48(6):691–727.

Mocanu, Delia, Andrea Baronchelli, Nicola Perra, Alessandro Vespignani, Bruno Goncalves & Qian Zhang. 2013. "The Twitter of Babel: Mapping World Languages through Microblogging Platforms." *PLOS One* 8(4):e61981.

Monroe, Burt L., Jennifer Pan, Margaret E. Roberts, Maya Sen & Betsy Sinclair. 2015. "No! Formal Theory, Causal Inference, and Big Data Are Not Contradictory Trends in Political Science." *PS: Political Science & Politics* 48(01):71–74.

Moore, Will H. 1995. "Rational Rebels: Overcoming the Free-Rider Problem." *Political Research Quarterly* 48(2):417–454.

Moore, Will H. 1998. "Repression and Dissent: Substitution, Context, and Timing." *American Journal of Political Science* 42(3):851–873.

Moore, Will H. 2000. "The Repression of Dissent: A Substitution Model of Government Coercion." *Journal of Conflict Resolution* 44(1):107–127.

Morstatter, Fred, Jurgen Pfeffer, Kathleen M. Carley & Huan Liu. 2013. Is the Sample Good Enough? Comparing Data from Twitters Streaming API with Twitters Firehose. In *Association for the Advancement of Artificial Intelligence*.

Mourtada, Racha & Fadi Salem. 2011. Civil Movements: The Impact of Facebook and Twitter. Technical Report 2 Dubai Scool of Government.

Muchnik, L., S. Aral & S. J. Taylor. 2013. "Social Influence Bias: A Randomized Experiment." *Science* 341(6146):647–651.

Mueller, Andreas. 2015. "scikit-learn.".
URL: *http://scikit-learn.org/stable/*

Mungiu-pippidi, Alina & Igor Munteanu. 2009. "Moldova's "Twitter Revolution"." *Journal of Democracy* 20(3):136–142.

Nardulli, Peter F., Scott L. Althaus & Matthew Hayes. N.d. "A Progressive Supervised-learning Approach to Generating Rich Civil Strife Data." *Social Methodology*. Forthcoming.

Newman, M.E.J. 2003. "The structure and Function of Complex Networks." *SIAM Review* 45(2):167–256.

Nielsen, Richard. 2012. "Jihadi Radicalization of Muslim Clerics.".

Opp, Karl-Dieter. 2012. "Collective identity, rationality and collective political action." *Rationality and Society* 24(1):73–105.

Opp, Karl-Dieter & Christiane Gern. 1993. "Dissident Groups, Personal Networks, and Spontaneous Cooperation: The East German Revolution of 1989." *American Sociological Review* 58(5):659–680.

Pappas, Takis S. 2008. "Political Leadership and the Emergence of Radical Mass Movements in Democracy." *Comparative Political Studies* 41(8):1117–1140.

Pei, Sen, Lev Muchnik, José S. Andrade Jr., Zhiming Zheng & Hernán A. Makse. 2014. "Searching for superspreaders of information in real-world social media." *Scientific Reports* 4(c):5547.

Pfaff, Steven. 1996. "Collective Identity and Informal Groups in Revolutionary Mobilization: East Germany in 1989." *Social Forces* 75(1):91–117.

Pierskalla, Jan H. & Florian M. Hollenbach. 2013. "Technology and Collective Action: The Effect of Cell Phone Coverage on Political Violence in Africa." *American Political Science Review* 107(02):207–224.

Pierskalla, Jan Henryk. 2010. "Protest, Deterrence, and Escalation: The Strategic Calculus of Government Repression." *Journal of Conflict Resolution* 54(1):117–145.

Press, Daryl G. 2001. "The Myth of Air Power in the Persian Gulf War and the Future of Warfare." *International Security* 26(2):5–44.

Quattrone, George a. & Amos Tversky. 1988. "Contrasting Rational and Psychological Analyses of Political Choice." *The American Political Science Review* 82(3):719.

Radsch, Courtney. 2008. "Core to Commonplace: The evolution of Egypts blogosphere." *Arab Media & Society* (6):1–14.

Radsch, Courtney C. 2012. "Unveiling the Revolutionaries: Cyberactivism and the Role of Women in the Arab Uprisings.".

Rahimi, Babak. 2011. "The Agonistic Social Media: Cyberspace in the Formation of Dissent and Consolidation of State Power in Postelection Iran." *The Communication Review* 14(3):158–178.

Raleigh, Clionadh, Andrew Linke, Havard Hegre & Joakim Karlsen. 2010. "Introducing ACLED: An Armed Conflict Location and Event Dataset: Special Data Feature." *Journal of Peace Research* 47(5):651–660.

Rasler, Karen. 1996. "Concessions, Repression, and Political Protest in the Iranian Revolution." *American Sociological Review* 61(1):132–152.

Reuter, Ora John & Graeme B. Robertson. 2015. "Legislatures, Cooptation, and Social Protest in Contemporary Authoritarian Regimes." *The Journal of Politics* 77(1):235–248.

Reza, Sadiq. 2007. "Endless Emergency: The Case of Egypt." *New Criminal Law Review* 10(4):532–553.

Riker, William H. & Peter C. Ordeshook. 1968. "A Theory of the Calculus of Voting." *The American Political Science Review* 62(1):25–42.

Rizzo, Helen, Anne M Price & Katherine Meyer. 2012. Targeting Cultural Change in Repressive Environments: The Campaign against Sexual Harassment in Egypt. Technical Report 614 Egyptian Center for Women's Rights Cairo: .
URL: *http://ecwronline.org/pdf/studies/AntiHarassment_for_ECWR.pdf*

Robertson, Graeme B. 2007. "Strikes and Labor Organization in Hybrid Regimes." *American Political Science Review* 101(04):781–798.

Rod, Espen Geelmuyden & Nils B. Weidmann. 2015. "Empowering activists or autocrats? The Internet in authoritarian regimes." *Journal of Peace Research* 52(3):338–351.

Romero, Daniel M, Brendan Meeder & Jon Kleinberg. 2011. Differences in the Mechanics of Information Diffusion Across Topics: Idioms, Political Hashtags, and Complex Contagion on Twitter. In *International World Wide Web Conference*. pp. 695–704.

Schelling, Thomas C. 1978. *Micromotives and Macrobehavior*. New York City: W.W. Norton & Company.

Schrodt, Philip A., Shannon G. Davis & Judith L. Weddle. 1994. "Political Science: KEDS–A Program for the Machine Coding of Event Data." *Social Science Computer Review* 12(4):561–587.

Schussman, Alan & Sarah Anne Soule. 2005. "Process and Protest: Accounting for Individual Protest Participation." *Social Forces* 84(2):1083–1108.

Seabold, Skipper & Josepf Perktold. 2014. "statstools.".
URL: *https://pypi.python.org/pypi/statsmodels*

Segerberg, Alexandra & W. Lance Bennett. 2011. "Social Media and the Organization of Collective Action: Using Twitter to Explore the Ecologies of Two Climate Change Protests." *The Communication Review* 14(3):197–215.

Shachar, Ron & Barry Nalebuff. 1999. "Follow the Leader: Theory and Evidence on Political Participation." *The American Economic Review* 89(3):525–547.

Shah, Dhavan V., Joseph N. Cappella & W. Russell Neuman. 2015. "Big Data, Digital Media, and Computational Social Science: Possibilities and Perils." *The ANNALS of the American Academy of Political and Social Science* 659(1):6–13.

Shapiro, Jacob N & David A Siegel. 2015. "Coordination and security: How mobile communications affect insurgency." *Journal of Peace Research* pp. 1–11.

Shapiro, Jacob N. & Nils B. Weidmann. 2011. "Talking About Killing: Cell Phones, Collective Action, and Insurgent Violence in Iraq.".

Siegel, David A. 2009. "Social Networks and Collective Action." *American Journal of Political Science* 53(1):122–138.

Siegel, David A. 2011. "When Does Repression Work? Collective Action in Social Networks." *The Journal of Politics* 73(04):993–1010.

Sinha, Shiladitya, Chris Dyer, Kevin Gimpel & Noah A Smith. 2013. "Predicting the NFL Using Twitter.".

Skocpol, Theda. 1979. *States & Social Revolutions: A Comparative Analysis of France, Russia, and China*. Cambridge: Cambridge University Press.

Snow, David A. & Dana M. Moss. 2014. "Protest on the Fly: Toward a Theory of Spontaneity in the Dynamics of Protest and Social Movements." *American Sociological Review* 79(6):1122–1143.

Snow, David A., Louis A. Zurcher Jr. & Sheldon Ekland-Olson. 1980. "Social Networks and Social Movements: A Microstructural Approach to Differential Recruitment." *American Sociological Review* 45(5):787–801.

Starbird, Kate & Ley Palen. 2010. Pass It On?: Retweeting in Mass Emergency. In *Information Systems for Crisis Response and Management*. Number December 2004 Seattle: pp. 1–10.

Starbird, Kate & Leysia Palen. 2012. (How) Will the Revolution be Retweeted? Information Diffusion and the 2011 Egyptian Uprising. In *CSCW*. Seattle: .

Stefanidis, Anthony, Andrew Crooks & Jacek Radzikowski. 2011. "Harvesting ambient geospatial information from social media feeds." *GeoJournal* 78(2):319–338.

Steinert-Threlkeld, Zachary C. 2016. "Spontaneous Collective Action: Peripheral Mobilization During the Arab Spring.".

Steinert-Threlkeld, Zachary C., Delia Mocanu, Alessandro Vespignani & James Fowler. 2015. "Online social networks and offline protest." *EPJ Data Science* 4(1):19.

Stone, Biz. 2010. "Tweet Preservation.".
URL: *https://blog.twitter.com/2010/tweet-preservation*

Stringhini, Gianluca, Gang Wang, Manuel Egele, Christopher Kruegel, Giovanni Vigna, Haitao Zheng & Ben Y. Zhao. 2013. Follow the Green: Growth and Dynamics in Twitter Follower Markets. In *Proceedings of the 2013 conference on Internet measurement conference*. pp. 163–176.

Suh, Bongwon, Lichan Hong, Peter Pirolli & Ed H. Chi. 2010. Want to be Retweeted? Large Scale Analytics on Factors Impacting Retweet in Twitter Network. In *IEEE Second International Conference on Social Computing*. pp. 177–184.

Taylor, Michael. 1988. Rationality and Revolutionary Collective Action. In *Rationality and Revolution*, ed. Michael Taylor. Cambridge: Cambridge University Press pp. 63–97.

Thelwall, Mike. 2008. "Social networks, gender, and friending: An analysis of MySpace member profiles." *Journal of the American Society for Information Science and Technology* 59(8):1321–1330.

Tonkin, Emma, Heather D. Pfeiffer & Greg Tourte. 2011. "Twitter, Information Sharing and the London Riots?" *Bulletin of the American Society for Information Science and Technology* 38(2):49–57.

Trejo, Guillermo. 2014. "The Ballot and the Street: An Electoral Theory of Social Protest in Autocracies." *Perspectives on Politics* 12(02):332–352.

Tucker, Joshua a. 2007. "Enough! Electoral Fraud, Collective Action Problems, and Post-Communist Colored Revolutions." *Perspectives on Politics* 5(03):535.

Tufekci, Zeynep. 2008. "Grooming, Gossip, Facebook and Myspace." *Information, Communication & Society* 11(4):544–564.

Tufekci, Zeynep. 2014. "Social Movements and Governments in the Digital Age: Evaluating a Complex Landscape." *Journal of International Affairs* 68(1):1–18.

Tufekci, Zeynep & Christopher Wilson. 2012. "Social Media and the Decision to Participate in Political Protest: Observations From Tahrir Square." *Journal of Communication* 62(2):363–379.

Tufekci, Zeynep & Deen Freelon. 2013. "Introduction to the Special Issue on New Media and Social Unrest." *American Behavioral Scientist* 57(7):843–847.

Tullock, Gordon. 1971. "The Paradox of Revolution." *Public Choice* 11:89–99.

Twitter. 2013. Annual Report 2013. Technical report.

Twitter. 2014. Selected Company Metrics and Financials. Technical report.

Ugander, Johan, Lars Backstrom, Cameron Marlow & Jon Kleinberg. 2012. "Structural diversity in social contagion." *PNAS* 109(16):5962–5966.

Update on the Twitter Archive At the Library of Congress. 2013. Technical Report January Library of Congress Washington D.C.: .

Valenzuela, Sebastian. 2013. "Unpacking the Use of Social Media for Protest Behavior: The Roles of Information, Opinion Expression, and Activism." *American Behavioral Scientist* 57(7):920–942.

Valkanas, George, Ioannis Katakis, Dimitrios Gunopulos & Antony Stefanidis. 2014. Mining Twitter Data with Resource Constraints. In *2014 International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)*. Ieee pp. 157–164.

Vieweg, Sarah, Amanda L. Hughes, Kate Starbird & Leysia Palen. 2010. Microblogging During Two Natural Hazards Events : What Twitter May Contribute to Situational Awareness. In *Human Factors in Computing Systems*. Atlanta: pp. 1079–1088.

Ward, Brian. 2014. "TCAT: The New Twitter Modeling Tool for Visualizing Social Media Data.".
URL: *http://painepublishing.com/measurementadvisor/pretty-pictures-new-twitter-modeling-tool-can-make-social-media-data-tangible-actionable/*

Ward, Michael D., Andreas Beger, Josh Cutler, Cassy Dorff & Ben Radford. 2013. "Comparing GDELT and ICEWS Event Data.".

Warren, T. Camber. 2014. "Not by the Sword Alone: Soft Power, Mass Media, and the Production of State Sovereignty." *International Organization* 68(01):111–141.

Watts, Duncan J. 2002. "A simple model of global cascades on random networks." *Proceedings of the National Academy of Sciences* 99(9):5766–5771.

Watts, Duncan J. 2004. "The New Science of Networks." *Annual Review of Sociology* 30(1):243–270.

Weyland, Kurt. 2012. "The Arab Spring: Why the Surprising Similarities with the Revolutionary Wave of 1848?" *Perspectives on Politics* 10(04):917–934.

Wooldridge, Jeffrey M. 2003. "Cluster-Sample Methods in Applied Econometrics." *American Economic Review* 93(2):133–138.

Wu, Shaomei, Jake M. Hofman, Winter a. Mason & Duncan J. Watts. 2011. "Who says what to whom on twitter." *Proceedings of the 20th international conference on World wide web - WWW '11* p. 705.

Xie, Wei, Cheng Li, Feida Zhu, Ee-Peng Lim & Xueqing Gong. 2012. When a friend in Twitter is a friend in life. In *Proceedings of the 4th Annual ACM Web Science Conference*. New York, New York, USA: ACM Press pp. 344–347.

Yardi, Sarita & Danah Boyd. 2010. Tweeting from the Town Square : Measuring Geographic Local Networks. In *Fourth International AAAI Conference on Weblogs and Social Media*. pp. 194–201.

Zamal, Faiyaz Al, Wendy Liu & Derek Ruths. 2012*a*. Homophily and Latent Attribute Inference: Inferring Latent Attributes of Twitter Users from Neighbors. In *Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media*. pp. 387–390.

Zamal, Faiyaz Al, Wendy Liu & Derek Ruths. 2012*b*. Homophily and Latent Attribute Inference: Inferring Latent Attributes of Twitter Users from Neighbors. In *Sixth International AAAI Conference on Weblogs and Social Media*. pp. 387–390.

Zeitzoff, T., J. Kelly & G. Lotan. 2015. "Using social media to measure foreign policy dynamics: An empirical analysis of the Iranian-Israeli confrontation (2012-13)." *Journal of Peace Research* 52(3):368–383.

Zheludev, Ilya, Robert Smith & Tomaso Aste. 2014. "When can social media lead financial markets?" *Scientific Reports* 4:4213.