Who Protests? Using Social Media Data to Estimate How Social Context Affects Political Behavior

Ted Hsuan Yun Chen* Paul Zachary Christopher J. Fariss

October 24, 2017

Abstract

Who protests? Prior scholarship on collective action identifies numerous individual-level causal mechanisms but measures these factors at the group-level. This is because individual-level data on participation or context are difficult to obtain. We address this inferential problem with data from the 2015 Freddie Gray protests in Baltimore, MD. We argue that arbitrary or capricious encounters with civil servants generate grievances. To assess the relationship between this grievance generating mechanism and the decision to protest, we obtain every tweet made during the protests and train a classifier model to estimate who protested. Next, we use a novel algorithm on users' tweet history to estimate where in the city users socialize. We estimate variation in anti-police grievances using these locations and geospatial arrest data from the Baltimore Police Department. Our results show that grievances — measured by exposure to policing events — are significantly correlated with an individual's participation in the protest.

*Ted Hsuan Yun Chen is a PhD candidate in the Department of Political Science at the Pennsylvania State University (thc126@psu.edu), Paul Zachary is a Visiting Fellow at the Institute for Quantitative Theory and Methods at Emory University (paul.zachary@emory.edu), and Christopher J. Fariss is Assistant Professor in the Department of Political Science, University of Michigan, Ann Arbor (cjfariss@umich.edu). We thank conference attendees at MPSA, PolNet, and PolMeth and workshop participants at Emory and the University of Washington for helpful feedback and suggestions. All errors are the authors'. The project has been reviewed and approved by the University of Michigan Institutional Review Board (HUM00126649), and determined not to be human research by the Pennsylvania State University Institutional Review Board (STUDY00007059). We acknowledge research support from the Charles Koch Foundation, Emory's Institute for Quantitative Theory and Methods, the National Science Foundation's XSEDE program (SES170013), and the San Diego Supercomputer Center.

1 Introduction

Some of the most memorable and consequential political events stem from protests. By aggregating disparate voices under a single banner, these collective action events have the potential to communicate the preferences of groups that political elites and lawmakers might not otherwise hear (Gause 2017; Davenport 2005; Williamson, Skocpol, and Coggin 2011; Chong 2014). Beyond their expressive value, collective action can change public opinion and influence policy (Madestam et al. 2013; Lohmann 1995). Despite their potential for effecting real policy change, no protest is ever universally attended. This is especially so at the start. For every person who joins in collective action events, others abstain and remain at home. These contextual variations beg the question: who protests?

Prior research on collective action gives unsatisfactory answers to this question. Rather, it offers important insights into related but different research questions such as the determinants of protest onset.[1] Existing formal models show that "tipping points" (Kuran 1989) or "informational cascades" (Lohmann 1994) can spur protests on. Protests begin when the risk of joining is less than citizens' costs of suppressing their anti-government feelings (Kuran 1989: 12-13). Participation then grows when attendance reveals information about social discontent to more moderate citizens (Lohmann 1994: 52-54).

How are these cascades triggered? These aforementioned formal models assume activists form a subset of the population, but do not explain why some groups' preferences are more extreme. Political context can affect behavior by generating grievances (Enos 2014). Scholars of relative depravation such as Gurr (1970) posit that perceived economic inequalities generate grievances in society. These horizontal inequalities among groups motivate participation in collective action in order to secure tangible economic or political benefits (Cederman, Wimmer, and Min 2010; Cederman, Weidmann, and Gleditsch 2011). As the grievances produced by inequality grow in intensity, elites aggregate demands into social movements. This suggests groups at the tails of the distribution of economic goods — either poor or rich — should be more likely to protest.

Despite its intuitive appeal, prior research finds little empirical support in the quantitative literature for the contention that grievances lead to collective action (Fearon and Laitin 2003). As qualitative accounts of collective action continue to stress the role of grievances in motivating participation, this missing empirical link is surprising (Wood 2003). One possible explanation is we measure grievances with "inappropriate conceptualization and imperfect measurements" Cederman, Weidmann, and Gleditsch (2011: 478). While scholars commonly measure grievances at the group-level, relative depravation theories suggest people derive their perceptions from their local context. Rather than rule out grievance as a motivation for participation in collective action, prior studies lack sufficiently granular data to test how perceptions affect behavior.

Though group-level factors influence a given groups' likelihood of participating in collective action, we agree that they are insufficient to explain individual variation. We argue that encounters with state bureaucracy and civil servants can generate individual-level grievances, and that these grievances make it more likely an individual participates in collective action. This is because civil servants, such as teachers, social workers, and law enforcement, provide important information

[1] There is also a well-developed literature on how leaders organize protests. See Walls (2015), Han (2014) and Ganz (2009).

about an individual's place in society. Civil servants, commonly referred to as "street-level bureaucrats," have lax oversight and wide discretion in their decision making (Lipsky 2010). This makes their decision-making capricious. Depending upon an individual's influence, street-level bureaucrats can use their discretion to facilitate or hinder their access to government resources. As few people have sufficient connections to influence bureaucrats, every encounter signals to citizens that the government is at best ambivalent or at worst abusive to their interests (Carrington 2005). This implies the more a citizen interacts with street-level bureaucrats, the more intense their grievances become.[2]

Substantial empirical evidence bolsters this last claim. Encounters with street-level bureaucrats decrease people's trust in government and perceived belonging in their communities (Lerman and Weaver 2014; Uggen and Manza 2002). In this paper, we contribute to the growing literature on the consequences for encountering the "carceral state." Over and above other street-level bureaucrats, police encounters have an especially strong and negative effect on trust in government and political engagement (Kang and Dawes 2017; Tyler, Fagan, and Geller 2014; Wildeman 2014; Uggen and Manza 2002).[3] This negative effect is not isolated to those directly interacting with law enforcement; it can spread through social networks (Burch 2013). Interactions with law enforcement also generates strong grievances due to perceptions of unfair targeting and second-class citizenship (Lerman and Weaver 2014). This suggests that people in high police activity ares should be more likely to engage in collective action related to reforming the police. As grievances are associated with particular institutions, this theory does not suggest increased participation in events that are substantively unrelated.

Grievances alone are an insufficient explanation of protest participation. In all but the rarest of cases, protests are planned and scheduled in advance (Han 2014; Payne 2007; Ganz 2009). Without knowing when and where a protest is scheduled, participation is impossible. We argue that increased demand for information is the mechanism through which grievances translate into participation. Our theory predicts community members with more intense grievances should have greater demand for information for logistical information from community organizers when grievance generating events take place (Zuern 2011; Norris, Walgrave, and Van Aelst 2005; Wood 2003; Murdie and Bhasin 2011).[4] As we do not expect people with low grievances to seek out information about logistics, they should be unlikely to appear at protests.

In most cases, social scientists lack individual-level data on participation and/or grievances. This forces quantitative scholars to use proxies for grievances such as access to foreign media or measures of economic inequality (Kern 2011; Steinert-Threlkeld et al. 2015; Steinert-Threlkeld 2017).[5] These measures, however, are either time-invariant or coarse aggregates. Without granular data, we cannot rule out grievance and perceptions as motivators for collective action.

[2]For them to effect, these grievances must be substantively related to the protest. Our theory does not predict that encounters with civil servants drives participation in all forms collective action.

[3]This decrease in trust occurs even when an interaction does not result in arrest.

[4]Although it is an important question, understanding activists' choice of when and where to hold protests is beyond this paper's scope.

[5]Other scholars such as Lawrence (2016) and Aytaç, Schiumerini, and Stokes (2017) survey protest participants to overcome this data limitation. Inferences from surveys present a number of methodological concerns such as desirability bias and homophily. We are unable to locate an example of a study that surveys both protest participants and abstainers, making it impossible to empirically assess the differences between these two groups.

To address this inferential problem, we introduce a new method that enables us to estimate individual-level social context via social media data. We study the protests that took place in Baltimore, MD in April 2015 against police brutality.[6] One issue when studying the police in the United States is the high correlation between a neighborhood's racial composition and police activity; police maintain a stronger presence and are more likely to use coercive authority in communities of color (Smith 1986). While parsing out the effect of race from police contact would be difficult in these localities, Baltimore exhibits uniquely significant variation in levels of police activity and race. This makes it an ideal test case for our theory.

Our research design is as follows. First, we first obtain every geotagged tweet made within the city in April 2015 and hand code a random subsample according to whether it indicates the person attended the Freddie Gray protests. Second, we use this training set to train a machine learning model to classify the remainder of our corpus of tweets. Third, we purchase the entire tweet history from accounts from our initial sample and then pass every user's tweets through a novel community detection algorithm. This new algorithm, based on earlier work by Rossmo (1999), uses the location where every tweet was posted to estimate where each user spends time throughout their day (Chen, Fariss, and Zachary 2017). These areas, which we refer to as social activity hubs (SAHs), allow us to estimate each user's political context. Finally, we use georeferenced arrest data from 2012-2015 that we obtained from the Baltimore Police Department through a public records request. Using these data, we find a robust and positive association between police contact and protest participation.

This paper makes several contributions to the theoretical literature on collective action. First, we show that qualitative accounts were right not to abandon their focus on grievances as motivators for participation. Our results, however, broaden prior studies' conceptualization of grievances to include political context. While material deprivation can produce grievances, our results suggest that the way the state interacts with its citizens affects their behavior. Second, we begin to investigate the microfoundations of compliance and enforcement. As prior studies focus on compliance among institutions, courts, and countries, we know relatively little about the mechanisms through which governments engender compliance with the law among their citizens (Carrubba 2005). Our results suggest additional avenues for research about the downstream consequences of enforcement on political behavior. Third, our results provide clarity about the social mechanisms that influence an individual to choose to join protests even when the costs are high — such as in autocratic societies.

Finally, we introduce a new estimator of social context from social media data. While social media is an increasingly common sampling frame, there was no way to measure where users reside in observational studies. This limits both the types and the quality of inferences that could be drawn from social media. Our estimator, which measures social context down to the neighborhood level, gives researchers an important new tool.

2 Who Protests?

Understanding collective action is one of the oldest research questions in social science (Gurr 1970; Olson 1965; Ostrom 1990; Popkin 1979; Tullock 1971). Prior research can be divided into two sep-

[6] As they took place after the death of Freddie Gray in police custody, these protests are commonly referred to as the Freddie Gray protests.

arate research questions. The first focuses on the link between economic depravation and collective action (Gurr 1970). They posit collective action is more likely in groups within the tails of the distribution of social goods. The second explores how opportunity structures affect participation (Earl et al. 2004; Tullock 1971; Schussman and Soule 2005; Lichbach 1994). In this section, we describe these prior studies' insights before noting there is no satisfactory explanation of individual-level variation in participation *within* groups. To address this gap, we propose a novel theory of how social context affects grievances, which in turn help explain who protests and when.

2.1 Relative Depravation and Inequality

While they highlight different causal mechanisms, relative depravation-based and inequality-based theories tie collective action to grievances produced by an individual's access to material goods (Gurr 1970, 1993; Cederman, Weidmann, and Gleditsch 2011). For relative depravation scholars, the primary source of social discontent is unfulfilled material aspirations. The first step in the causal process is for an individual to perceive a gap between the standard of living they believe they deserve and what they actually have. Taking stock of this gap creates grievances, which interpersonal and inter-group comparisons then exacerbate. As this perceived gap grows, moreover, it produces more intense grievances. As Regan and Norton (2005: 320-321) summarize, the "causal mechanism is the psychological process in which an individual compares their current situation against their expected standard of living." These perceptions translate into collective action once elites aggregate channels these feelings into a social movement.

Although both focused on standard of living, a different causal mechanism generates grievances from inequality. If relative depravation focuses on perceived inequalities, then inequality highlights real ones. Grievances begin once an individual compares their level of material wealth to others in society. These grievances grow more intense as the gap in material wealth grows between the individual and their reference group. Inequality does not exclusively generate grievances among the poor. As Cederman, Weidmann, and Gleditsch (2011) and Boix (2008) argue, economic inequality can also incentivize the rich to mobilize to prevent economic transfers to the poor. Economic inequalities translate into collective action after elites aggregate demands for social and economic justice into a social movement.

While theories of relative depravation and inequality generally link the distribution of material wealth to grievances, wealth is not necessarily the only such source. Grievances come from unequal access to whatever the relevant economic cleavage in society is. Prior scholarship identifies sources of grievances as diverse as access to land rents (Thomson 2016; Acemoglu and Robinson 2006); misallocation of investment (Robinson and Torvik 2005); and natural resource rents (Ross 2004). Moreover, the inequality need not be economic. Perceived status inequalities can also trigger collective action as groups use protests to reposition themselves in the social hierarchy or demand more rights (Horowitz 1985; Chong 2014; Willer 2009).

Regardless of the precise source of the inequality, these theories observable implication is that collective action is more likely in communities along economic extremes than those closer to the median.

2.2 Opportunity Structures

Even if relative depravation and inequality are intuitively appealing, there is very mixed empirical support that links the distribution of material goods to collective action (Fearon and Laitin 2003; Hegre and Sambanis 2006; Lichbach 1989). Moreover, relative depravation loses some of its theoretical appeal under closer inspection. As Snyder and Tilly (1972) note, every society is replete with aggrieved, angry, and frustrated people. Especially in poor communities, real and perceived inequality and depravation is near-universal. Yet, protest and other forms of collective action remain rare.

If material depravation is universal and time-invariant, it is insufficient to explain participation in collective action. In their most reduced form, theories of political opportunity structure suggest collective action when "the political environment that provide[s] incentives for people to undertake collective action" changes such that it affects potential participants' "expectations for success and failure" (Tarrow 1994: 85). Such changes can include exogenous developments such as technological development (Steinert-Threlkeld 2017) or endogenous weakening of state repressive institutions (Fearon and Laitin 2003). In all cases, collective action is determined by citizens ability to organize and the states capacity to manage opposition.

The first set of political determinants of contentious politics are factors that influence citizens willingness to organize protest and realize potential risks from collective action. Citizens considering joining in protests encounter challenges from both the state and others in their group (Kuran 1989; Lohmann 1994). The state can use violence and repression to diminish turnout (Davenport 2010; Davenport, Soule, and Armstrong 2011; Lorentzen 2014). Protestors confront collective action problems in both the presence or absence of repression (Olson 1965). Even when protests to secure policy redress is Pareto-improving — when the change in allocation over the desired good makes one or more individuals better off without making another worse off — each individual should rationally abstain and free-ride off of others. This suggests that a time-varying mechanism enables citizens to organize and overcome this collective action problem.

What separates first movers from other citizens in their ability to solve collective action problems? The literature identifies two answers to this question. The first is that some citizens — such as activists and community organizers — might receive especially high amounts of utility from political engagement. This increased utility makes activists "extraordinarily willing to take risks for a worthy political cause" (Lawrence 2016: 4). While the literature is unclear about the exact source of this increased utility, it might simply derive from the "pleasure of agency" (Wood 2003: xv-xvi). The other is that first movers are unusually interested in acquiring some selective incentive such as political power or resources (Popkin 1988; Schneider and Teske 1992). Another mechanism that solves the collective action problem for first movers is selective incentives (Lichbach 1994, 1998; Mason 1984).

2.3 Grievances and the Selective Incentive for Information

In this section we devise an argument that links information seeking and grievances. Individuals interact with state security agents in a variety of settings. Sometimes these interactions are without incident. Under other circumstances however, grievances are generated by the interaction. If well-known means to address the grievance exist then information seeking takes places through established institutional channels (e.g. through the judicial system). However, if such institutions

do not exist or are not trusted then information seeking must, by necessity, take place outside of established institutional channels. This type of information seeking behavior creates a selective incentive to organize by the individuals directly affected by an interaction that generates the grievance. When the aggrieved do not trust the courts or the police, then organizing a group to seek information about the event by protest becomes an important alternative option. These first movers solve the collective action problem for themselves because they have a selective incentive to organize. Beyond this, as these individuals organize into a larger groups the costs of joining in the protest decreases, which therefore solves the collective action problem for others.

The collective action problem exists because the costs of participating in action against the state are high and if the movement is successful all members of the group enjoy the benefits regardless of whether they participated in the movement or not (Lichbach 1994, 1998; Mason 1984; Olson 1965). In order to solve collective action problems selective incentives may be used to compel or even coerce participation (Lichbach 1994, 1998; Mason and Krane 1989). Selective incentives are private or club goods available to individuals participants in the collective action but not every member of the group. Selective incentives can include wages, paid now or in the future, or other incentives obtainable by the individual through participation in the group.[7]

Building on prior research on collective action, we argue that individuals with more protest-relevant grievances should be more willing to participate in collective action because they have a selective incentive to obtain information — an instrumental motive. Instrumental motivation theories suggest that horizontal inequalities should largely determine behavior (Gurr 1970; Cederman, Weidmann, and Gleditsch 2011). Once activists overcome collective action problems, members of social classes who are deprived relative to the reference group join in. While this helps explain why some classes are — on average — more likely to mobilize than others, it does not explain individual-level variation *within* these groups.

In this section, we develop a novel theory that explains why some join protests while others abstain within the same social grouping. Our theory highlights the role played every citizen's social and political context in producing grievances (Enos 2016; Enos and Gidron 2016; Enos 2014; Sands 2017). Even within the same community, citizens might have widely divergent lived experiences. Some community members might be satisfied with the status quo, while others might be highly motivated to seek political redress.

The literature on conflict onset highlights a number of mechanisms that produce grievances. Horizontal inequalities, or "inequalities in economic, social or political dimensions or cultural status between culturally defined groups," are one such mechanism (Stewart 2008: 3). These inequalities might arise when a group perceives it does not have its fair share of society's wealth (Collier and Hoeffler 2004). Inversely, poor economic conditions can also make groups more likely to mobilize

[7]Lichbach (1994) argues that collective actions problems are solvable when selective incentives are used; however, selective incentives applied without reference to a collective goal, often represented by an ideology, do not result in successful movements. Specifically, Lichbach (1994) argues that the collective action problem is more likely to be solved under three conditions: (1) private goods are available as benefits, (2) the costs and benefits applied to each peasant is different given the level of participation, (3) the distribution rule that governs access to the good is contingent on participation. According to Lichbach (1994), collectively supplied benefits are "public goods", which are nontrivial and excludable. "The term thus covers objectives that are both reformist (e.g., having the existing government supply a new road) and revolutionary (e.g., taking over the government and supplying your own road" (Lichbach1994: 388, note 25).

in order to protect the current distribution of goods (Miguel, Satyanath, and Sergenti 2004).

Demands for prestige and status can also mobilize groups into collective action (Sambanis 2001; Cederman, Wimmer, and Min 2010). The literature on ethnicity and civil conflict frequently links nationalism and demands for self-determination with mobilization (Fearon, Kasara, and Laitin 2007; Buhaug, Cederman, and Rød 2008). Laws and institutions that exclude specific groups from power serve as a focal point for organizing and solving collective action problems. This dynamic can be seen in the United States from the Civil Rights-era protests as well as the ongoing LGBT equality movement (Chong 2014; Davenport 2007).

These prior theories intentionally "[shift] the explanatory focus from individualist to group-level accounts" of mobilization (Cederman, Weidmann, and Gleditsch 2011: 477). While group-level accounts help explain why — on average — some groups are more likely to mobilize than others, they cannot explain variation in participation *within* groups. While we concur with theories focusing on horizontal inequalities, we focus instead on how context affects how individuals perceive their position within their cultural group. In our empirical work below, we explore how interactions with bureaucrats affect grievances.

As Lipsky (2010) argues, the quality and nature of an individual's interactions with government employees are one important signal of their context and position within their group. These "street-level bureaucrats," such as social workers, police officers, or school teachers, wield enormous power over citizens' lives. They determine whether a permit is issued; a child is removed from their parent's custody; or a fine is imposed. Importantly, these bureaucrats have tremendous discretion and as such can determine how much access a citizen has to government services. As citizens use these interactions to update upon their support for the government, every interaction with a street-level bureaucrat has the possibility of generating new grievances (Lipsky 2010; Butler and Broockman 2011; White, Nathan, and Faller 2015).

Although our theory generalizes to interactions with all bureaucrats, our empirical focus in this paper is the effect of encounters with police officers on grievances and mobilization. Interacting with police officers is rarely pleasant and people of color must be especially vigilant to protect their physical integrity (Davenport 2005, 2010). This observation is highlighted by the growing literature on the consequences of interactions with the "carceral state," which shows that police-citizen interactions decrease trust in government and participation in politics through formal institutions (Kang and Dawes 2017; Tyler, Fagan, and Geller 2014; Lerman and Weaver 2014; Wildeman 2014; Uggen and Manza 2002). These negative effects also propagate locally within social networks to those who did not interact directly with the police (Burch 2013).

This observation suggests that police encounters produce grievances at the individual-level. over and above those horizontal inequalities generate, these additional grievances from social interactions explain who joins during collective action events.

How do these grievances help individuals overcome collective action problems? We argue that grievances affect demand for information about protests and opportunities for collective action. Logistical information such as the time and date of a planned protest is necessary for participation. Although we do not test this mechanism directly, we expect grievances to increase information-seeking behavior through social media or personal social networks.

Acquiring this information is important because it is inefficient for organizers to try to mobilize people who prefer the status quo. Second, after identifying aggrieved community members, they use these individuals to disseminate information within their social networks. Because organizers

cannot talk to everyone in the community, they rely upon the people they contact to spread information through their social networks. This exposes individuals with higher levels of grievance to community organizers and activists (Zuern 2011; Norris, Walgrave, and Van Aelst 2005; Wood 2003; Murdie and Bhasin 2011).

3 Research Design

Studying collective action presents numerous measurement problems. While most theories make individual-level predictions about behavior, measures of participation and/or grievances are rarely available at such fine resolution.[8] To surmount this problem, scholars typically rely on aggregate measures such as access to foreign media or self-reports of behavior in surveys (Kern 2011; Steinert-Threlkeld et al. 2015; Steinert-Threlkeld 2017; Lawrence 2016; Aytaç, Schiumerini, and Stokes 2017). Each approach limits the types of behavioral inferences that can be made. Inferring individual behavior from aggregate data can produce highly inaccurate, biased estimates (King 1997). Known variously as aggregation bias or the ecological inference problem, this statistical property makes inferences about individual-level behavior derived from aggregate data particularly fragile.

While they do not suffer from aggregation bias, surveying protestors poses different challenges. First, social desirability bias likely influences respondents' answers in ways that are correlated with their social context and the protest's outcome. A respondent might overstate their participation after successful protests or diminish it after failed ones. Second, it is incredibly difficult to validate the representativeness of a sampling frame. This issue is non-trivial because scholarship on homophily suggests that participants will segregate within protests (Clack, Dixon, and Tredoux 2005).

To overcome these challenges, scholars have recently turned to social media data to measure patterns of protest and mobilization (Steinert-Threlkeld et al. 2015; Steinert-Threlkeld 2017; Lawrence 2016). While social media data has the advantage of permitting scholars to measure protestors' social networks and explore how information spreads within networks, it has the drawback of lacking the contextual and demographic information that comes with surveys. Twitter does not disclose its users race, age, or gender and offers only rudimentary tools to infer location.[9] This forces scholars to use aggregate data, subject to the bias described above.

By introducing a novel method to estimate social context from social media data, we propose a novel solution to these inferential problems. We test for how grievances influence protest participation during a highly contentious and violent series of protest and repression events that occurred during April 2015 in Baltimore, MD following the death of Freddie Gray in Baltimore Police Department (BDP) custody. We focus on Baltimore because it fits the criteria of a "typical" case, in which "cases are intended to represent descriptive features of a broader set of cases" (Gerring 2008). Typical cases are particularly beneficial for theory testing. The correlation between community demographics and police activity is very strong in the United States (Smith 1986). Unlike other

[8] As Cederman, Weidmann, and Gleditsch (2011: 478) note, the "formidable problems of data availability associated with the uneven coverage and comparability of" individual-level measures of grievances and/or mobilization, "most scholars have had to content themselves with selective case studies or statistical samples restricted to particular world regions."

[9] Researchers have access to each account's time zone and language setting.

American cities, Baltimore has numerous communities of color with high and low police activity. Without this variation, it would be impossible to parcel the effect of race out from that of police contact.

BPD officers arrested Gray on April 12, 2015 for suspected possession of an "illegal switchblade."[10] Gray fell into a coma for reasons that are disputed while being transported in a police van. He did not recover from his injuries and died on April 19. To denounce Gray's alleged mistreatment and BDP brutality, protestors began gathering in front of the Western District Police Station on April 18. These protests grew rapidly in size throughout the week and eventually reached several thousand people and only ended after the Maryland National Guard restored calm to the city by imposing a mandatory curfew from April 28 to May 3. These protests, collectively referred to as the Freddie Gray protests, are our empirical focus.

Our research design is as follows. First, we purchased every geotagged tweet made from within Baltimore, MD from April 16, 2015 to May 4, 2015, i.e. during the Freddie Gray protests.[11] This generated a total of 111,440 tweets from 7,884 unique users. Second, we hand coded 5,000 randomly selected tweets as to whether each tweet indicated participation in the protests.[12] Third, hand coded tweets serve as a training set for a classifier model that we use to classify our remaining corpus of tweets as to whether they indicate protest participation. Our results suggest that approximately 5% of our corpus of tweets indicate protest participation.

Third, we purchase the entire account history from users in our data. Fourth, after obtaining the account histories, each account is passed through a novel community detection algorithm (Chen, Fariss, and Zachary 2017). Based on earlier work on community detection by Rossmo (1999), our algorithm identifies clusters in locations were users tweet in order to estimate their movement patterns. These areas, which we call social activity hubs (SAHs), enable us to estimate users' political context as measured by where they spend time. Finally, we measure grievances with georeferenced arrest data we obtained from from the Baltimore Police Department through a public records request. Using these data, we find a robust and positive association between police contact and protest participation.

3.1 Protester Classification

In order to develop a classifier model of protest-related tweets, we manually coded a random subsample of 5,000 tweets drawn from our full corpus. We assign tweets a 1 when it indicates the user was physically present at the Freddie Gray protests and assign tweets a 0 when they are unrelated to the protests.[13] For examples of both protest and non-protest related tweets, see Figure 1:

[10] While officers testified they believed Gray's knife was illegal, the Maryland State's Attorney for Baltimore later clarified that he was in fact in possession of a "spring-assisted knife" that was legal under Maryland law (Blinder and na 2015).

[11] Location sharing is an opt-in feature and is disabled by default. One concern is that users who opt into sharing this data are systematically different than those who do not. Addressing this question, Pavalanathan and Eisenstein (2015) find that users who enable geolocation are demographically similar to other Twitter users.

[12] Two research assistants read and coded the subsample separately. The intercoder reliability was over 99%.

[13] 624 tweets, or approximately 12%, of our training set suggested the user attended the protests.

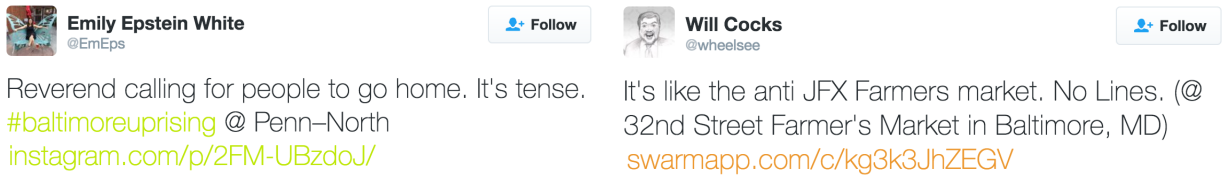

Figure 1: Left: A protest-related tweet referencing ongoing discussions about whether protestors would comply with the curfew order. **Right:** A non-protest related tweet.

3.1.1 Preprocessing

Because Twitter imposes a limit of 140 characters on all tweets, each tweet's brevity poses unique inference problems (Saif et al. 2014; Naveed et al. 2011).Users frequently abbreviate or use slang to maximize the content of each tweet. This implies there can be substantial variation in language use patterns, even among Twitter users discussing the same topic. Such variation, coupled with the brevity of each tweet, results in highly-dimensional, sparse data. As there may not be overlap in the context of tweets among users, this variation poses problems for text analysis algorithms. Insufficient overlap is a problem, because it means our training set will be uninformative, resulting in low accuracy scores.

To address this problem, we preprocess our data extensively prior to analysis. First, we use the CamelCase package in Python to split hashtags into separate words whenever possible.[14] Second, we increase the textual overlap in our corpus by stemming and normalizing our corpus with NLTK's TweetTokenizer package. By stemming our corpus, we reduce inflected words to their word stem.[15] In order to normalize our corpus and reduce dimensionality, we convert emoji into unicode; preserve punctuation; and remove accents from letters.

Finally, after normalizing our corpus, we tokenize our hand coded training set and obtain feature vectors. We obtain features using term-frequency inverse-document frequency (TF-IDF). While other tokenization algorithms weigh every feature's importance by counting the number of times it appears in a corpus, TF-IDF improves upon this approach accounting for both the number of times each feature appears and the total number of words in the corpus. We allow our $n-$gram size to range from 1 to 5 and converted to binary with a threshold of 0. We then convert these features into a vector in which each element represents a single feature (obtained via TF-IDF) and its value is its weight. In order words, we transform every tweet into a vector in which each unique word is an element and that element's value is the frequency that word appears in our corpus divided by the weight.

3.1.2 Classification

We estimate the probability the tweets outside of our training set indicate the poster attended the Freddie Gray protests in person. After preprocessing, we randomly divide our manually coded tweets into a training (80%) and a test (20%) set for cross-validation. We fit two models after

[14]For instance, the common hashtag #FreddieGrayProtest became "Freddie Gray protest" after processing.

[15]A word stem is the grammatically uninflected root word. For example, "paste" is the stem of words such as "pastes," "pasted," and "pasting."

performing a grid search: regularized logistic regression (LR) and a support vector classifier (SVC) with feature selection.[16]

To estimate the probability each tweet belongs to a protestor, LR models fit a linear regression using all features as regressors. The output is then the estimate of the probability with which each observation belongs to either category using the logistic function. SVCs identify the minimal support vectors, which distinguish texts in one class versus another. Specifically, we collapse our training data into a matrix ($\mathbf{X}$), where each item x_{ij} indicates the j features used by Twitter user i. As we coded whether each Twitter user protested (p_i,) we can write our training data as (x_{ij}, p_i). The SVC attempts to separate this data such that we can describe the best hyperplane whereby one set of Twitter users has $p_i = 1$ while the other has $p_i = 0$..[17]

SVCs identify a vector (w) such that w characterizes the two hyperplanes that maximize the separation between the two classes, i.e. the w whereby $w \cdot x_i - b \geq 1$ is true. In other words, SVCs minimize $\|w\|$ such that $\forall_i = 1, \ldots, n p_i(w \cdot x_i - b) \geq 1$. The SVC classifies the remaining corpus of tweets by projecting them onto the hyperplane generated previously. This projection is then converted into a probability distribution after fitting a sigmoid (Platt 1999).

As our SVC and LR both performed well, we averaged their predictions for our remaining corpus using soft voting, which sums the predicted probabilities across an ensemble of classifiers. We then evaluate the quality of our predictions with their Matthews correlation coefficient.[18] After crossvalidation on our test set, we estimate that the final accuracy of our model using soft voting to be 95.12% ($\sigma = 0.50$) with a false positive rate of 13.92% ($\sigma = 3.53$).

In order to maximize the size of each user's corpus of tweets, we aggregate our final estimation up to the account-level. We then fit our final model onto the entire uncoded set. We code a user as attending the Freddie Gray protest if at least one of their tweets is protest-related. At the user level, we simulating classification errors to estimate our false positive rate: 8.21% ($\sigma = 1.80$). Figure 2 shows the proportion of users and tweets that our classification procedure estimates indicate protest participation:

3.2 Social Activity Hubs

Although the above enables us to identify which users attended the Freddie Gray protests, it does not provide us with information about the users' social and political context. We argue that context conditions political behaviors such as protest participation. More explicitly, this implies that exposure to social and political cues present within the environments affects political behavior. Testing this argument requires some method to place users in geographic space. In this section,

[16] Although naive Bayes is a common text classification algorithm, our initial grid research suggested its accuracy for our particular application was significantly lower than other approaches.

[17] A hyperplane is the set of points in vector x such that $w \cdot x - b = 0$.

[18] Matthews correlation coefficients (MCC) are a correlation coefficient between the observed and predicted binary classification (Matthews 1975). As a measure of the quality of binary classifications, MCCs are particularly robust to instances when the two classes are of very different sizes. MCCs can be calculated from the confusion matrix as: $MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN(FN+FP)(TN+FN)}}$, where TP is the number of true positives; TN the number of true negatives; FP the number of false positives; and FN the number of false negatives.

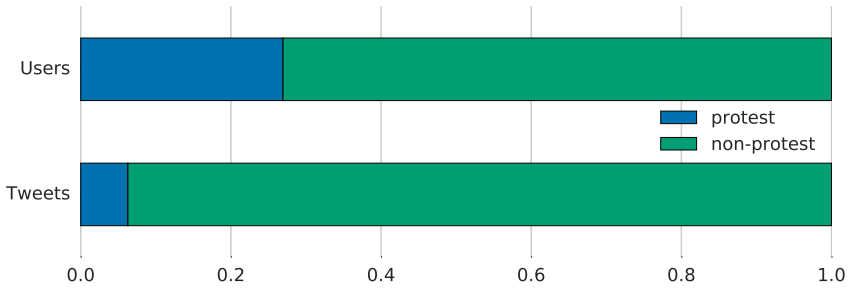

Figure 2: This bar plot shows the proportion of tweets and users that indicate protest participation. The proportion of users classified as participating in the protest exceeds the proportion of tweets discussing the protest because protestors used Twitter to discuss both the protest as well as other subjects.

we describe our novel method that estimates Twitter users' "social activity hubs" (SAHs), or the centroid of the areas where users spend time throughout the day (Chen, Fariss, and Zachary 2017).

Prior to estimating users' SAHs, we collected each user's historical tweets using Twitter's API, which we called through the TwitteR package in R (Gentry 2015).[19] Each geotagged tweet is treated as an observed incident of the user's movement patterns, and from the collection of all observed incidents, we estimate the user's SAHs. As the availability of information associated with each Twitter account differs, our SAH model, summarized in algorithm 1, is conditional on what this information affords, defaulting to more basic models where data availability is low. More specifically, we define SAHs in two ways described in more detail below, based on posterior quantities of a Dirichlet process mixture (DPM) model for spatial data (Verity et al. 2014). Where these estimates are unavailable due to convergence failure in the MCMC (implemented in the Rgeoprofile package for R, Stevenson et al. 2014), we take the incident closest in Euclidean distance to the spatial mean of all observed incidents as the SAH. For users with only one observed incident ($n = 197$), that incident is taken as the SAH.

In the remainder of this section, we describe the DPM model in more detail, discuss how we use its posterior quantities in our SAH model, and explain how we account for uncertainty in the model. We conclude by presenting some ethical considerations.

3.2.1 Dirichlet Process Mixture Model for Spatial Data

For user whose tweets contain enough information regarding their movement patterns, we use the Dirichlet process mixture (DPM) model of geographic profiling as the basis of our SAH model. DPM models for spatial data, based on prior geographic profiling models in criminology (O'Leary 2010; Rossmo 1999), was first described in (Verity et al. 2014) where it was applied to spatial epidemiology. More recently the model was applied in an attempt to determine the identity of graffiti artist Banksy (Hauge et al. 2016). The intuition of the DPM model for spatial data is to sort a set of observed incidents in physical space into clusters originating from different source locations without prior assumptions about the number of clusters that exist. For our present purposes, the DPM model is preferred over alternatives that require a fixed number of clusters (including those

[19]Twitter only allows access to users' 3,200 most recent tweets through its API. Tweets were collected between July 19, 2016 and August 27, 2016. API Documentation available here: `https://dev.twitter.com/rest/reference/get/statuses/user_timeline`.

with a single cluster) because individuals are likely to vary in terms of their movement patterns for we have no prior information. Where there are multiple clusters, especially when they are highly dispersed, a misspecified number will result in inaccurate source location estimates skewed by "outliers" which are actually observations originating from another source. The DPM model rectifies this by estimating the number of sources based on the observed data. The flexibility afforded by this feature is especially desirable given the large number of Twitter accounts we are working with, as it is not feasible to tweak the SAH model for each Twitter account individually. A DPM model is not without assumptions, which are apparent from the technical description below. In short, by employing the DPM model, we are assuming that individuals can have multiple SAHs from where their movement outward follows identical distributions, which in this implementation we specify as a bivariate normal distribution with standard deviation of approximately two miles.[20]

More explicitly, the DPM model we use, adapted for spatial data by Verity et al. (2014), is as follows. For each Twitter user, define a two-dimensional sample space with a finite grid of cells as Ω, in which each cell $\boldsymbol{\omega} = (\omega^{(1)}, \omega^{(2)})$ is a vector containing the latitude and longitude in decimal degrees of a geocoordinate. The set of n geocoordinates obtained from geotagged tweets $\mathbf{x} = \mathbf{x}_{1,\ldots,n}$ is assumed to be the result of independent draws from a mixture of a countably infinite set of bivariate normal distributions centered on $\mathbf{z} = \mathbf{z}_{1,\ldots,\infty}$, each with a variance of σ^2 ($\sigma = 0.05$). Both $\mathbf{x}$ and $\mathbf{z}$ are defined on Ω. The prior distribution of the set of $\mathbf{z}$ is assumed to be a bivariate normal centered on the mean of $\mathbf{x}$, with a variance of τ^2 (τ is set to the largest distance in either longitude or latitude). c_i is a categorical variable that assigns $\mathbf{x}_i$ to source $\mathbf{z}_{c_i}$, and is drawn from a Dirichlet process, specifically the Chinese Restaurant Process which has a concentration parameter α drawn from a diffuse hyper-prior (specifically $h(\alpha) = ((1+\alpha)^2)^{-1}$) and a base distribution that is the bivariate normal (with mean $\mathbf{x}/n$) discussed above. The above is formally represented as,

$$\begin{aligned} \mathbf{x}_i | \mathbf{z}_{c_i} &\sim \mathcal{N}(\mathbf{z}_{c_i}, \Sigma = \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix}) \\ \mathbf{z}_{1,\ldots,\infty} &\sim \mathcal{N}(\mathbf{x}/n, \mathrm{T} = \begin{bmatrix} \tau^2 & 0 \\ 0 & \tau^2 \end{bmatrix}) \\ c_i &\sim CRP(\alpha) \\ \alpha &\sim \mathcal{H} \end{aligned} \tag{1}$$

Exact computation of posterior quantities are intractable when the number of observations is high ($n > 10$ being a useful rule of thumb; see Verity et al. 2014: for analytical solutions to relevant posterior quantities), but can be estimated using MCMC methods (Neal 2000; Verity et al. 2014), which is implemented in the R package Rgeoprofile 1.2 (Stevenson et al. 2014). The MCMC algorithm (`RunMCMC()` presented in algorithm 2) is based on a Gibbs sampler that alternates between draws of of source locations $\mathbf{z}_{c_i}$ and cluster assignment c_i for all $i = 1, \ldots, n$ observations. The algorithm returns, for each $\mathbf{x}_i$, its cluster c_i; and for each unique cluster c_j, its spatial mean $\mathbf{z}_j$.

3.2.2 Local Minima and Cluster Mean Submodels

As introduced earlier, we use the posterior quantities obtained from the DPM model in our SAH model in two ways. For the *local minima* submodel, begin by defining $\mathbf{S} \subseteq \Omega$ as the grid bound by the minimum and maximum values of the set of observed $\mathbf{x}$. Next, for every cell $\mathbf{s} \in \mathbf{S}$, rank $\mathbf{s}$ according to the sum of its distances to each source location $\mathbf{z}_j$ over all posterior draws, where

[20] Specifically 0.05 decimal degrees which translates to approximately 2 miles in Baltimore, MD.

distance is not linear but weighed by the inverse of the bivariate normal density around $\mathbf{z}_j$. In concordance with existing geoprofiling approaches (e.g. Rossmo 1999), ranks are transformed to hit scores on $[0, 1)$, but remain functionally equivalent in that lower is better and all values are distinct.[21] The type of hit score surface is traditionally used as a surface for search priority of source locations (Rossmo 1999; Verity et al. 2014). On this surface, we find all m local minima (i.e. locations with higher priority) within an approximately two mile radius (0.05 decimal degrees) and define a user's SAHs as the set of m observed $\mathbf{x}_i$ closest to these local minima. For the *cluster mean* submodel, we define a user's SAHs as the set of observed $\mathbf{x}_i$ closest to the set of estimated source locations $\mathbf{z}$ averaged across posterior draws.

Based on each user's geoprofile, we take the most likely sources, up to three locations, and set the incidents (i.e. geocoded tweets) closest to them as the user's SAH, and the census tract associated with that location as the user's SAH neighborhood. Not all estimated SAHs are within Baltimore City bounds, but given the constraint that we only have arrest data from the Baltimore Police Department, described later, we dropped all SAHs not in Baltimore. Results from the DPM model for all users are presented in Figure 3. For users with multiple estimated SAHs within Baltimore, we take the average of the census tract characteristics across all estimated neighborhoods, but note that given the parameters of the DPM model (i.e. the standard deviation of the bivariate normal distribution in the migration profile), it is rare for users to have more than two estimated SAHs within Baltimore.

Figure 4 illustrates the SAHs estimated under both models in relation to the hit score surface produced by the DPM model.

3.2.3 Uncertainty in DPM Model

In earlier applications of the DPM model to spatial data, there is justifiably less of a concern over the uncertainty of estimates. However uncertain, the expected values of $\mathbf{z}$ is what informs a search that must take place. Existing implementation of the model (Stevenson et al. 2014) therefore do not readily yield uncertainty measures. For inferential modeling, however, measures of uncertainty feature much more prominently. In order to account for uncertainty in our SAH model, we take 2000 draws from the posterior distribution of the DPM model, thinned to 100 samples, and use this information to determine a set of corresponding SAHs following both the local minima and cluster mean submodels. Specifically, for the local minima submodel, instead of computing a hit score surface based on all posterior draws, we do so for each draw independently; and for the cluster mean submodel, $\mathbf{z}$ is not averaged across posterior draws. SAH estimates are stored for each posterior draw, forming a posterior distribution of SAHs. This distribution can be used in subsequent statistical modeling to account for uncertainty associated with the SAH model. In remainder of this section, we illustrate using a specific Twitter account what the uncertainty about the two DPM-based SAH submodels look like. This particular account was chosen because it is illustrative. The level of uncertainty associated with this account, based on visual inspection, is neither particularly high nor low.

In order to visualize the level of uncertainty about our SAH estimates, we plot the variation in hit scores associated with each potential source across all posterior draws. Future work might benefit from a formal quantification of the type of uncertainty discussed here. Specifically, in Figure 5

[21] The two computational steps above are implemented in the `ThinandAnalyse()` function in Rgeoprofile 1.2 (Stevenson et al. 2014).

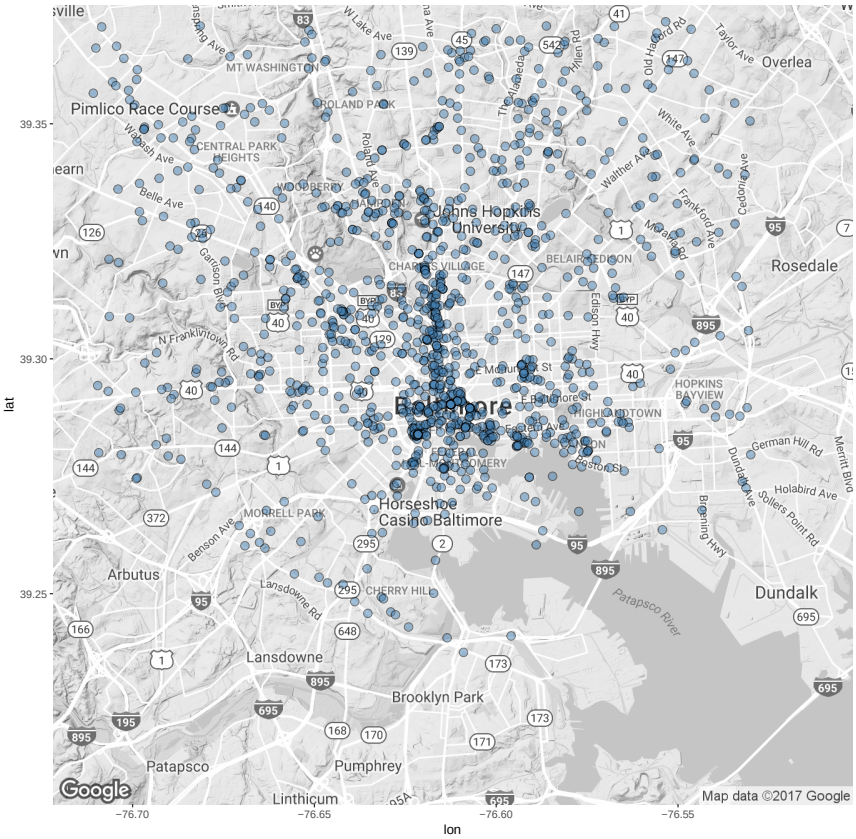

Figure 3: Geographical distribution of social activity hubs estimated using the DPM model.

(which corresponds to the hit score surface in Figure 4), each horizontal line documents the hit score of a particular source location as it varies across posterior draws. The highlighted lines are the sources chosen as the SAHs from the combined posterior draws, which may differ depending on the model used. Variation in the hit score indicates changes in the topography of the hit score map across posterior draws. This does not, however, necessarily mean there is uncertainty about the SAH estimates, which arise when changes in the topography are large enough to induce changes in the hit score rank of potential sources relative to each other, indicated by crossing lines.

In this particular case, both methods led to the potential source location with the lowest hitscore as one of the two sources. In this particular example, we see that one source (the one highlighted in blue) has very little uncertainty about it, while the other (in red) has higher uncertainty. The blue line is not straight, meaning that the hitscore map topography around it changes from draw to draw, but it remains consistently low and never crosses with another line, meaning that it is always the source with the best hitscore across all draws.

3.2.4 Ethical Considerations

As our research strategy enables us to estimate Twitter users' social activity hubs, we take several steps to maintain anonymity and protect users from potential harm. First, our sample is only taken from users who had opted into sharing their location with Twitter. By default, Twitter does not record the location where a tweet was posted. Instead, users must change their phone's settings to

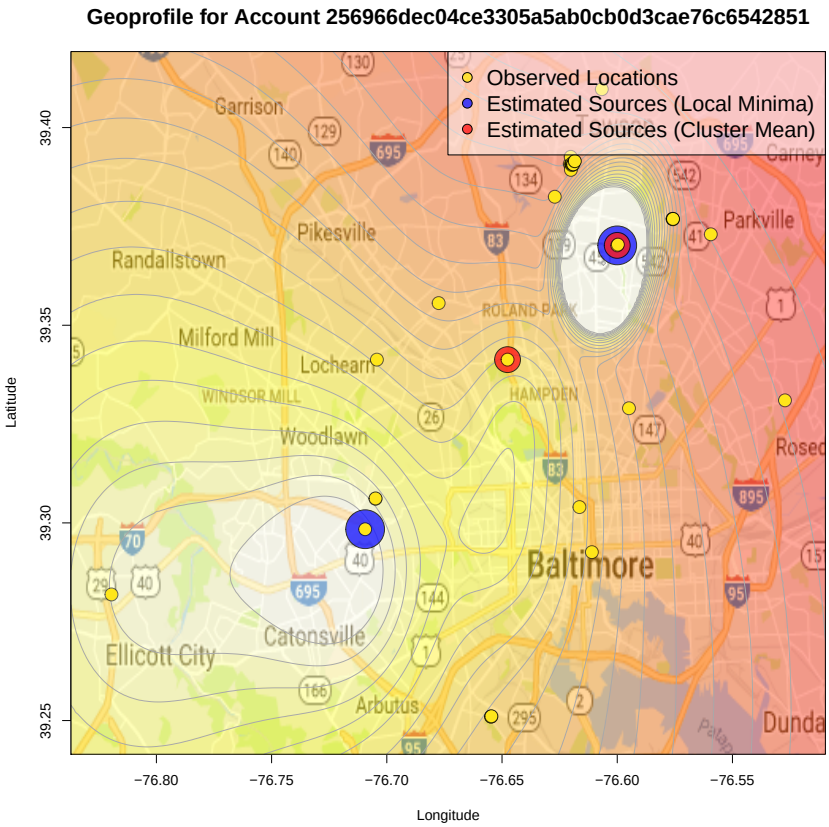

Figure 4: Example of a user's estimated social activity hub locations in relation to the hit score surface produced by the DPM model. Yellow points are observed incidents. Points enclosed in blue indicate SAHs determined by the local minima submodel. Points in enclosed in red indicate those determined by the cluster means submodel.

give Twitter permission to record their location via GPS. Second, we anonymize Twitter account names by applying a cryptographical hash. Hashing account names prevents anyone with access to either the original or replication data to identify and/or locate the users in our study.

3.3 Grievance Measurement

The Freddie Gray protests were directed against the Baltimore Police Department (BPD), which was perceived as hostile to the city's African-American community in the wake of widely publicized cases of police brutality. We proxy for the concentration of anti-BPD grievances using the location of arrests made by the BPD.

These data were obtained through a long-running public records request to the BPD. The data, released to us in 2017, contains records from the years 2012-2015 inclusive. In our initial request, we asked for copies of every arrest's narrative and/or arrest super form in order to code the amount of force used in each arrest. As these reports are not digitized by the BPD, releasing them would require years of redactions and organizing. As such, our request to access these reports was denied by the Chief of the BPD's Legal Affairs Division as "unduly burdensome."

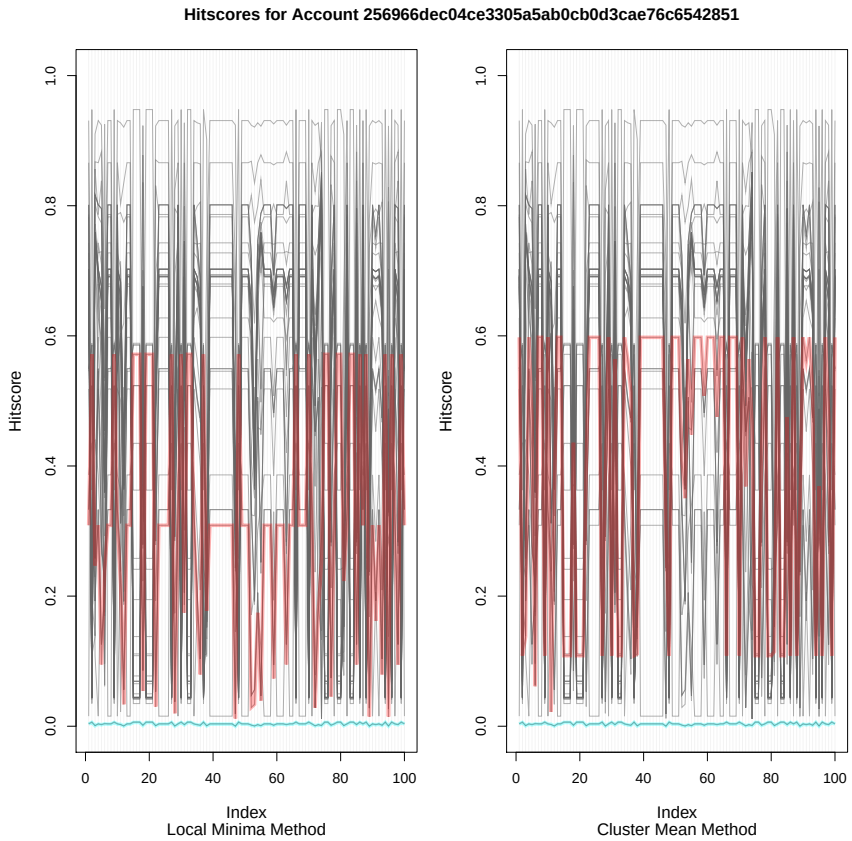

Figure 5: Example of a user's estimated social activity hub locations in relation to the geoprofile produced by the DPM model. Light blue points are observed incidents. Points enclosed in red indicate social activity locations under the DPM model and those enclosed in blue indicate points from the SAH neighborhood under the modal-tweet model.

Instead, the BPD provided arrest logs for all felony arrests. These logs contain the date the arrest was made; the location where the arrest was made; and the charges. Any personal information about the arrestee was redacted by the BPD under various subsections of the Maryland Public Information Act related to the release of personally-identifiable information.

We then convert the street address provided with each arrest to latitude and longitude via Google Map's public API. We use data from the 34,880 felony arrests made by the Baltimore Police Department from January 1, 2012 to April 17, 2015. Figure 6 shows the location of all unique visits by the BPD that resulted in at least one arrest between 2012 and 2014.

We use the arrest data to construct two measures of anti-BPD grievances, which is a result of exposure to policing. The first measure, *Total Arrests*, is simply the number of arrests within a census tract. The second measure, *Visits*, treats all arrests at a given location at the same time as a single incident within the census tract. Figure 7 shows the density of both measures of exposure to policing in Baltimore census tracts.

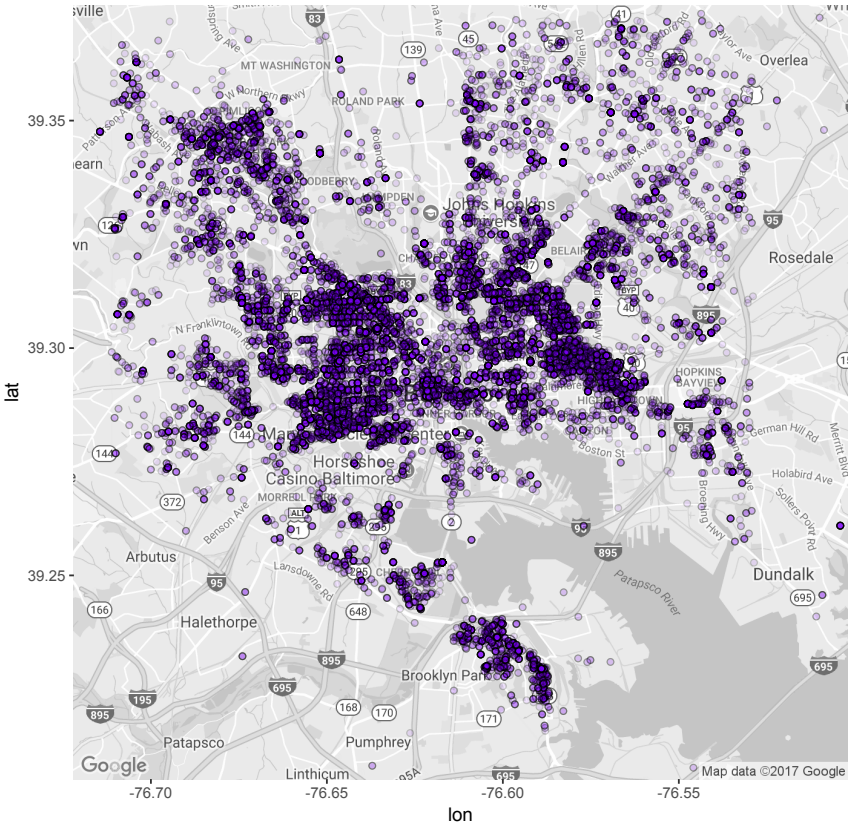

Figure 6: Geographical distribution of all unique visits resulting in one or more arrests by the BPD in 2012 and 2015. A small jitter (0.0005 to 0.001 degree decimals) is added to the coordinates.

3.4 Control Variables

In order to control for characteristics of each user's home neighborhood, we access information about the racial and demographic makeup of each Baltimore census tract as provided by the U.S. Census Bureau. Specifically, we include in our model the proportion of the population that is (1) Black, (2) Hispanic, (3) female, (4) between the ages of 15 and 24 (i.e. the youth bulge), (5) at least a high school graduate; (5) the median income, the (6) poverty and the (7) unemployment rates; and (8) the overall population. The distributions for these control variables are presented in Appendix Figure 11.

Our final model is as follows, with two types of exposure to policing as described above.

$$\begin{aligned} Protest \sim\ & Exposure\ \ to\ \ Policing + Proportion\ \ Black + Proportion\ \ Hispanic \\ & + Proportion\ \ Female + Youth\ \ Bulge + Median\ \ Income + Poverty\ \ Rate \\ & + Unemployment\ \ Rate + Proportion\ \ Highschool\ \ Graduate \\ & + Population + Population^2 + \epsilon \end{aligned} \tag{2}$$

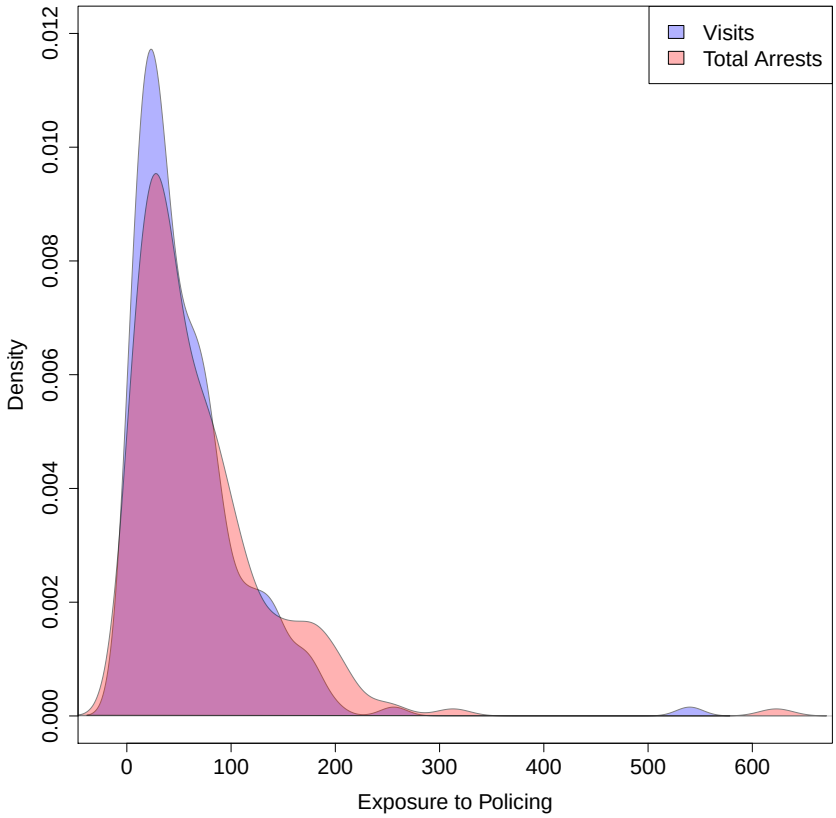

Figure 7: Density of both measures of exposure to policing in Baltimore City census tracts.

4 Results

Results from our models are presented in Table 1 and Table A1. The substantive effect of exposure to policing is graphically presented in Figure 8.

To understand the substantive impact of an additional year of tenure on homicide, we estimate the marginal effect of an additional year of tenure. After setting the other explanatory variables in Equation 1 to their mean, the predicted participation rate increases from 26% ($\sigma = 0.03$) for individuals experiencing the lowest level of police contact to 0.38% ($\sigma = 0.10$) experiencing the highest. These results are plotted in Figure 8.

Table 1: Models of Protest Participation Using Social Activity Hub

	Model I		Model II	
	Coef.	S.E.	Coef.	S.E.
Exposure to Policing†	0.10*	0.04	0.12*	0.05
Proportion Black	0.12	0.31	0.12	0.31
Proportion Hispanic	−2.59	1.79	−2.58	1.79
Proportion Female	−0.60	1.62	−0.60	1.62
Youth Bulge	0.73	0.83	0.73	0.83
Median Income	−1.06	0.92	−1.06	0.92
Poverty Rate	0.00	0.00	0.00	0.00
Unemployment Rate	0.14	0.74	0.15	0.74
Proportion HS Graduate	0.02	1.23	0.00	1.23
Per 1000 Population	0.22	0.25	0.22	0.25
Per 1000 Population^2	−0.05	0.04	−0.05	0.04
Intercept	−0.10	1.51	−0.09	1.51
N	1,239		1,239	
AIC	1478.3		1478.3	

* $p < 0.05$
† Exposure to policing is per 100 arrests in Model I and per 100 visits in Model II.

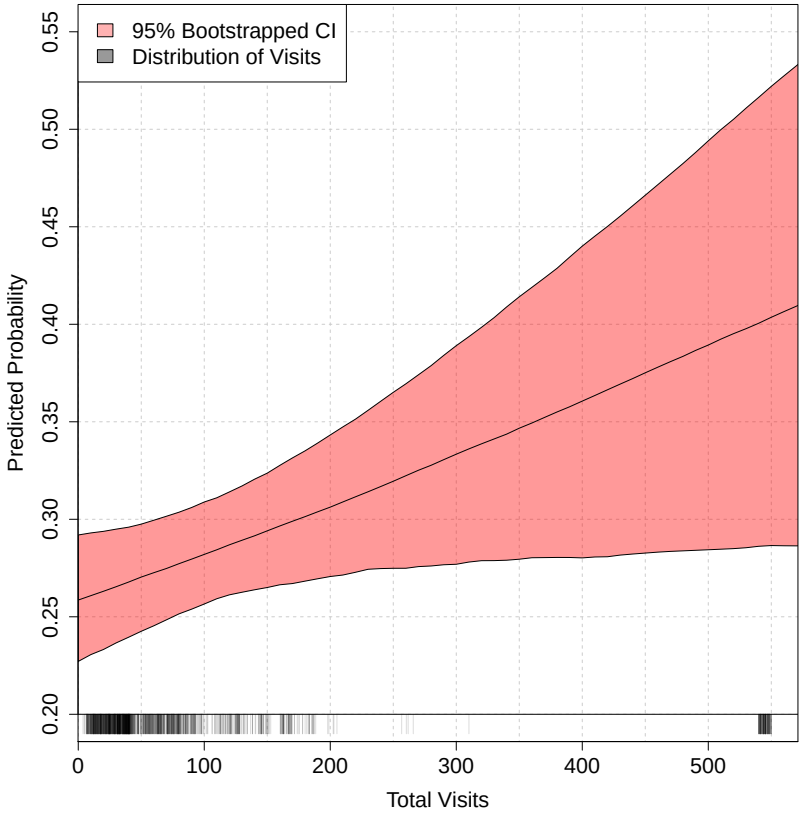

Figure 8: Predicted probability of being classified as a protest account simulated using estimated parameters from Model II (i.e. exposure to police is measured by total visits and home locations are estimated using the DPM). 95% confidence interval is obtained using the bootstrap method.

4.1 Mechanism

Our theory suggests that encounters with street-level bureaucrats produce grievances, which make participation in collective action more likely. Not all arrests produce grievances in the community. Lipsky (2010) argues that civil servants produce grievances when they exhibit discretion and preferential treatment for some group. In terms of arrests, this observation suggests that the only arrests that generate grievances are those where officers could have shown discretion and leniency but chose not to. This is similar to the negative effect on perceptions of government generated by traffic stops and stop-and-frisk, which rely on officers' subjective judgments (Kang and Dawes 2017; Lerman and Weaver 2014). We do not expect arrests for crimes where officers cannot show discretion to produce grievances.

In our data, grievance-generating arrests include those for marijuana and other drugs, because only a small portion of either drug users and sellers are ever arrested. In contrast, we do not expect arrests for murder or gun crime to generate grievances in the community because most agree severe crime deserves to be punished.

In Table 3, we subset our data on arrests according to what the BPD recorded as the initial charge. While the specifics of the charging document might change in between the time the person is first arraigned and is brought to trial, it is a useful proxy of the type of crime in which the arrestee was engaged. We use the 2017 Maryland Sentencing Guidelines Offense Table to code each crime according to its maximum term; offense type; and whether it involves marijuana or guns. Table 3 shows our estimated effect of drug arrests on participation in the Freddie Gray protests.

Table 2: Models of Protest Participation Using Social Activity Hub: Drug Arrests

	All Drugs				Marijuana			
	Model I		Model II		Model I		Model II	
	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Exposure to Policing†	0.19	0.08	0.21	0.09	2.85	0.90	3.39	1.07
Proportion Black	0.03	0.30	0.04	0.30	-0.01	0.30	-0.01	0.30
Proportion Hispanic	-2.99	1.73	-2.98	1.73	-2.70	1.75	-2.73	1.75
Proportion Female	-0.36	1.56	-0.39	1.56	-0.25	1.55	-0.27	1.55
Youth Bulge	0.43	0.82	0.44	0.82	0.20	0.83	0.22	0.83
Unemployment Rate	-1.35	0.91	-1.35	0.91	-1.59	0.92	-1.58	0.92
Median Income	-0.00	0.00	-0.00	0.00	-0.00	0.00	-0.00	0.00
Poverty Rate	-0.05	0.73	-0.03	0.73	0.01	0.73	0.02	0.73
Proportion HS Graduate	-0.02	1.19	-0.08	1.19	-0.11	1.19	-0.17	1.19
Per 1000 Population	0.23	0.25	0.24	0.25	0.16	0.25	0.16	0.25
Per 1000 Population^2	-0.05	0.03	-0.05	0.03	-0.04	0.03	-0.04	0.03
Intercept	0.09	1.48	0.15	1.48	0.39	1.49	0.42	1.49

† Exposure to policing is per 100 arrests in Model I and per 100 visits in Model II.

As predicted, we find a strong positive association between discretionary arrests such as arrests for

drugs and marijuana in a community with participation in collective action. In contrast, Table **??** the opposite. Examining arrests for murder and arrests with a maximum charge of life in prison, we find no association with protest participation.

Table 3: Models of Protest Participation Using Social Activity Hub: Non-Discretionary Arrests

	Murder				Life Sentence			
	Model I		Model II		Model I		Model II	
	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Exposure to Policing†	1.14	0.57	1.34	0.70	0.44	0.42	0.48	0.48
Proportion Black	-0.33	0.27	-0.33	0.27	0.04	0.30	0.04	0.30
Proportion Hispanic	-2.76	1.42	-2.76	1.42	-3.10	1.75	-3.11	1.75
Proportion Female	0.41	1.51	0.39	1.51	-0.75	1.59	-0.76	1.60
Youth Bulge	-0.27	0.75	-0.25	0.75	0.81	0.80	0.82	0.80
Unemployment Rate	0.98	0.82	1.00	0.82	-1.14	0.90	-1.13	0.90
Median Income	-0.00	0.00	-0.00	0.00	-0.00	0.00	-0.00	0.00
Poverty Rate	-0.26	0.71	-0.25	0.71	0.13	0.73	0.13	0.73
Proportion HS Graduate	-1.48	1.13	-1.47	1.13	-0.14	1.24	-0.13	1.25
Per 1000 Population	0.07	0.21	0.07	0.21	0.35	0.25	0.36	0.25
Per 1000 Population^2	-0.02	0.03	-0.02	0.03	-0.07	0.03	-0.07	0.03
Intercept	0.35	1.38	0.33	1.38	0.11	1.48	0.10	1.48

† Exposure to policing is per 100 arrests in Model I and per 100 visits in Model II.

4.2 Robustness

The results in Table 1 show that there is a strong positive association between encounters with the police and participation in collective action. In the appendix, we report alternative specifications as robustness tests. Appendix Table A1 confirms our results are not an artifact of our novel measurement of political context: social activity hubs. Rather than allowing a user's social activity to range across Baltimore, we measure each user's context using the location from which they tweet most frequently. Using this alternate measure does not substantially change our estimates of police contact's effect.

Table A2 considers whether our evidence for our mechanism is a product of our focus on particular crimes. Above, we show that arrests for murder or those that carry a maximum of a life sentence have no effect on participation in collective action. We posit this is because they do not generate grievances, as community members demand arrests for such crimes. In the appendix, we show similar evidence that arrests for "personal" crimes such as kidnapping and assault as well as crimes where the perpetrator was armed has no effect on collective action.

Another concern is the statistical significance of our measures of contact is model dependent. If that were the case, altering any of covariates on the right-hand side of our estimating equation would

alter or eliminate the observed statistical relationship. To assess this possibility, we estimate a least absolute shrinkage and selection operator (Lasso) regression to identify which variables minimize prediction errors. Both measures of contact with police remain important predictor variables. These results, presented in Appendix Table **??**, are strong evidence that our results are not model dependent.

Finally, we estimate whether there are any outliers in our data and whether they bias our results. To do so, we estimate the Cook's distance and DFBetas for our model. We find little evidence of outliers in our data. This suggests that our results are not driven by influential observations or outliers.

5 Discussion

The theory presented in this paper predicts that variation in grievance intensity can help explain participation in collective action. We argue that political context is an important — and overlooked — source of grievances (Enos 2014). Prior theoretical accounts of the link between grievance and costly political participation focuses on horizontal inequalities (Gurr 1970; Cederman, Weidmann, and Gleditsch 2011). Even if inequality between groups is informative regarding *which* groups mobilize, it does not explain why some people participate while others abstain. Building on Lipsky (2010), we contend that interactions with "street-level bureaucrats" is both contextually-dependent and an important source of grievances.

Data limitations have hitherto limited our ability to test conjectures about individual-level variation in grievances or protest participation; measuring whether an individual participated in a protest or their relevant grievances is non-trivial. Prior studies address this issue by either surveying protestors in the field (Lawrence 2016; Aytaç, Schiumerini, and Stokes 2017) or using observational designs (Madestam et al. 2013; Kern 2011). While field surveys yield measures participation, desirability bias might influence respondents' answers. Individual-level inferences from aggregate data in observational studies, moreover, might be biased.

We argue that it is essential to measure context at the individual-level in order to understand participation in collective action. To do so, we exploit social media data from the protests during April 2015 in Baltimore, MD after the death of Freddie Gray in police custody. After Gray was arrested, he fell into a coma in a police van and later died from his injuries. This triggered a wave of protests against police brutality throughout the city that only ended after the Maryland National Guard imposed a curfew. While participation was widespread, it was not complete. Drawing on the literature on the effect of encountering the carceral state, we argue that interacting with the police generate more intense grievances (White 2015; Uggen and Manza 2002). In turn, individuals subject to more police activity were more likely to protest.

We surmount earlier data limitations by purchasing every geotagged tweet made within Baltimore in April and May 2015. We then train a classifier model to identify whether every account belongs to a protestor. Despite sparse data, our classifier model has an accuracy rate of approximately 95%. We then purchase the users' entire Twitter history and pass the geotagged tweets through a novel community detection algorithm. Within the data, each community is a different area in Baltimore where the user spends time (Rossmo 1999; Verity et al. 2014). These communities, which we refer to as social activity hubs (SAHs), are a novel estimation strategy to measure context from social media data.

We find a strong positive relationship between police activity and participation in collective action. This finding has a number of implications for our understanding of political behavior and highlights opportunities for additional research. First, prior research on the carceral state strongly suggest interactions with police officers demobilize voters (Bruch, Ferree, and Soss 2010; Hjalmarsson and Lopez 2010; Burch 2011; Meredith and Morse 2015; Gerber et al. 2015; White 2015). Our results suggest police encounters have a displacing effect on political participation. Rather than decrease total levels of participation, police activity might instead displace participation into informal channels such as activism or protest. Linking social media accounts to voter files would be useful to test this conjecture.

References

Acemoglu, Daron, and James A Robinson. 2006. "Economic backwardness in political perspective." *American Political Science Review* 100 (1): 115–131.

Aytaç, S Erdem, Luis Schiumerini, and Susan Stokes. 2017. "Why Do People Join Backlash Protests? Lessons from Turkey." *Journal of Conflict Resolution* .

Blinder, Alan, and Richard Pérez-Pe na. 2015. "6 Baltimore Police Officers Charged in Freddie Gray Death." *New York Times* .

Boix, Carles. 2008. "Economic roots of civil wars and revolutions in the contemporary world." *World Politics* 60 (3): 390–437.

Bruch, Sarah K, Myra Marx Ferree, and Joe Soss. 2010. "From policy to polity: Democracy, paternalism, and the incorporation of disadvantaged citizens." *American Sociological Review* 75 (2): 205–226.

Buhaug, Halvard, Lars-Erik Cederman, and Jan Ketil Rød. 2008. "Disaggregating ethno-nationalist civil wars: A dyadic test of exclusion theory." *International Organization* 62 (3): 531–551.

Burch, Traci. 2011. "Turnout and party registration among criminal offenders in the 2008 general election." *Law & Society Review* 45 (3): 699–730.

Burch, Traci. 2013. *Trading democracy for justice: Criminal convictions and the decline of neighborhood political participation*. Chicago, IL: University of Chicago Press.

Butler, Daniel M, and David E Broockman. 2011. "Do politicians racially discriminate against constituents? A field experiment on state legislators." *American Journal of Political Science* 55 (3): 463–477.

Carrington, Keith. 2005. "Is there a need for control?" *Public Administration Quarterly* pp. 140–161.

Carrubba, Clifford J. 2005. "Courts and compliance in international regulatory regimes." *Journal of Politics* 67 (3): 669–689.

Cederman, Lars-Erik, Andreas Wimmer, and Brian Min. 2010. "Why do ethnic groups rebel? New data and analysis." *World Politics* 62 (1): 87–119.

Cederman, Lars-Erik, Nils B Weidmann, and Kristian Skrede Gleditsch. 2011. "Horizontal inequalities and ethnonationalist civil war: A global comparison." *American Political Science Review* 105 (3): 478–495.

Chen, Ted, Chris Fariss, and Paul Zachary. 2017. "Social Activity Hubs (SAHs): A New Method for Estimating User Specific Contextual Factors From Social Media Data. In *Computational Social Science 2017*, ed. David Krakauer, and Scott Page.

Chong, Dennis. 2014. *Collective action and the civil rights movement.* Chicago, IL: University of Chicago Press.

Clack, B., J. Dixon, and C. Tredoux. 2005. "Eating together apart: Patterns of segregation in a multi-ethnic cafeteria." *Journal of Community and Applied Social Psychology* 15: 1–16.

Collier, Paul, and Anke Hoeffler. 2004. "Greed and grievance in civil war." *Oxford economic papers* 56 (4): 563–595.

Davenport, Christian. 2005. "Understanding Covert Repressive Action: The Case of the U.S. Government against the Republic of New Africa." *Journal of Conflict Resolution* 49 (1): 120 – 140.

Davenport, Christian. 2007. "State repression and political order." *Annual Review of Political Science* 10: 1–23.

Davenport, Christian. 2010. *Media Bias, Perspective, and State Repression: The Black Panther Party.* Cambridge, MA: Cambridge University Press.

Davenport, Christian, Sarah A. Soule, and David A. Armstrong. 2011. "Protesting While Black? The Differential Policing of American Activism, 1960 to 1990." *American Sociological Review* 76 (1): 152–178.

Earl, Jennifer, Andrew Martin, John D. McCarthy, and Sarah A. Soule. 2004. "The use of newspaper data in the study of collective action." *Annual Review of Sociology* 30: 65–80.

Enos, Ryan D. 2014. "Causal effect of intergroup contact on exclusionary attitudes." *Proceedings of the National Academy of Sciences* 111 (10): 3699–3704.

Enos, Ryan D. 2016. "What the demolition of public housing teaches us about the impact of racial threat on political behavior." *American Journal of Political Science* 60 (1): 123–142.

Enos, Ryan D, and Noam Gidron. 2016. "Intergroup behavioral strategies as contextually determined: Experimental evidence from Israel." *The Journal of Politics* 78 (3): 851–867.

Fearon, James D, and David D Laitin. 2003. "Ethnicity, insurgency, and civil war." *American political science review* 97 (1): 75–90.

Fearon, James D, Kimuli Kasara, and David D Laitin. 2007. "Ethnic minority rule and civil war onset." *American Political science review* 101 (1): 187–193.

Ganz, Marshall. 2009. *Why David sometimes wins: Leadership, organization, and strategy in the California farm worker movement.* New York, NY: Oxford University Press.

Gause, LaGina. 2017. "The Advantage of Disadvantage: Legislative Responsiveness to Collective Action by the Politically Marginalized." Unpublished book manuscript.

Gentry, Jeff. 2015. *twitteR: R Based Twitter Client.* R package version 1.1.9.
URL: *http://CRAN.R-project.org/package=twitteR*

Gerber, Alan S, Gregory A Huber, Marc Meredith, Daniel R Biggers, and David J Hendry. 2015. "Can incarcerated felons be (Re) integrated into the political system? Results from a field experiment." *American Journal of Political Science* 59 (4): 912–926.

Gerring, John. 2008. "Case Selection for Case-Study Analysis: Qualitative and Quantitative Techniques." In *The Oxford handbook of political methodology*, ed. Janet M. Box-Steffensmeier, Henry E. Brady, and David Collier. New York, NY: Oxford University Press.

Gurr, Ted Robert. 1970. *Why Men Rebel*. Princeton, NJ: Princeton University Press.

Gurr, Ted Robert. 1993. "Why minorities rebel: A global analysis of communal mobilization and conflict since 1945." *International Political Science Review* 14 (2): 161–201.

Han, Hahrie. 2014. *How organizations develop activists: Civic associations and leadership in the 21st century*. New York, NY: Oxford University Press.

Hauge, Michelle V, Mark D Stevenson, D Kim Rossmo, and Steven C Le Comber. 2016. "Tagging Banksy: Using geographic profiling to investigate a modern art mystery." *Journal of Spatial Science* 61 (1): 185–190.

Hegre, Håvard, and Nicholas Sambanis. 2006. "Sensitivity analysis of empirical results on civil war onset." *Journal of conflict resolution* 50 (4): 508–535.

Hjalmarsson, Randi, and Mark Lopez. 2010. "The voting behavior of young disenfranchised felons: Would they vote if they could?" *American Law and Economics Review* pp. 356–393.

Horowitz, Donald L. 1985. *Ethnic groups in conflict*. Berkeley, CA: Univ of California Press.

Kang, Woo Chang, and Christopher Dawes. 2017. "The Electoral Effect of Stop-and-Frisk." SSRN Working Paper.

Kern, Holger Lutz. 2011. "Foreign media and protest diffusion in authoritarian regimes: The case of the 1989 East German revolution." *Comparative Political Studies* 44 (9): 1179–1205.

King, Gary. 1997. *A Solution to the Ecological Inference Problem: Reconstructing Individual Behavior from Aggregate Data*. Princeton, NJ: Princeton University Press.

Kuran, Timur. 1989. "Sparks and prairie fires: A theory of unanticipated political revolution." *Public choice* 61 (1): 41–74.

Lawrence, Adria K. 2016. "Repression and Activism among the Arab Spring's First Movers: Evidence from Morocco's February 20th Movement." *British Journal of Political Science* .

Lerman, Amy E, and Vesla M Weaver. 2014. *Arresting citizenship: The democratic consequences of American crime control*. Chicago, IL: University of Chicago Press.

Lichbach, Mark. 1998. *The Rebel's Dilemma*. Ann Arbor, MI: University of Michigan Press.

Lichbach, Mark I. 1994. "What Makes Rational Peasants Revolutionary? Dilemma, Paradox, and Irony in Peasant Collective Action." *World Politics* 46 (3): 383–418.

Lichbach, Mark Irving. 1989. "An evaluation of "does economic inequality breed political conflict?" studies." *World politics* 41 (4): 431–470.

Lipsky, Michael. 2010. *Street-level bureaucracy: dilemmas of the individual in public service*. New York, NY: Russell Sage Foundation.

Lohmann, Susanne. 1994. "The dynamics of informational cascades: The Monday demonstrations in Leipzig, East Germany, 1989–91." *World politics* 47 (1): 42–101.

Lohmann, Susanne. 1995. "Information, access, and contributions: A signaling model of lobbying." *Public Choice* 85 (3): 267–284.

Lorentzen, Peter. 2014. "China's Strategic Censorship." *American Journal of Political Science* 58 (2): 402–414.

Madestam, Andreas, Daniel Shoag, Stan Veuger, and David Yanagizawa-Drott. 2013. "Do political protests matter? evidence from the tea party movement." *The Quarterly Journal of Economics* p. qjt021.

Mason, T. David. 1984. "Individual Participation in Collective Racial Violence: A Rational Choice Synthesis." *American Political Science Review* 78 (4): 1040–1056.

Mason, T. David, and Dale A. Krane. 1989. "The Political Economy of Death Squads: Toward a Theory of the Impact of State-Sanctioned Terror." *International Studies Quarterly* 33 (2): 175–198.

Matthews, Brian W. 1975. "Comparison of the predicted and observed secondary structure of T4 phage lysozyme." *Biochimica et Biophysica Acta (BBA)-Protein Structure* 405 (2): 442–451.

Meredith, Marc, and Michael Morse. 2015. "The politics of the restoration of ex-felon voting rights: The case of Iowa." *Quarterly Journal of Political Science* 10.

Miguel, Edward, Shanker Satyanath, and Ernest Sergenti. 2004. "Economic shocks and civil conflict: An instrumental variables approach." *Journal of political Economy* 112 (4): 725–753.

Murdie, Amanda, and Tavishi Bhasin. 2011. "Aiding and abetting: Human rights INGOs and domestic protest." *Journal of Conflict Resolution* 55 (2): 163–191.

Naveed, Nasir, Thomas Gottron, Jérôme Kunegis, and Arifah Che Alhadi. 2011. Bad news travel fast: A content-based analysis of interestingness on twitter. In *Proceedings of the 3rd International Web Science Conference*. ACM p. 8.

Neal, Radford M. 2000. "Markov chain sampling methods for Dirichlet process mixture models." *Journal of computational and graphical statistics* 9 (2): 249–265.

Norris, Pippa, Stefaan Walgrave, and Peter Van Aelst. 2005. "Who demonstrates? Antistate rebels, conventional participants, or everyone?" *Comparative politics* pp. 189–205.

O'Leary, Mike. 2010. Implementing a Bayesian approach to criminal geographic profiling. In *COM. Geo*.

Olson, Mancur. 1965. *The Logic of Collective Action*. Cambridge, MA: Cambridge University Press.

Ostrom, Elinor. 1990. *Governing the Commons*. Cambridge, MA: Cambridge Cambridge University Press.

Pavalanathan, Umashanthi, and Jacob Eisenstein. 2015. Confounds and Consequences in Geotagged Twitter Data. In *Proceedings of Empirical Methods for Natural Language Processing (EMNLP)*.

Payne, Charles M. 2007. *I've got the light of freedom: The organizing tradition and the Mississippi freedom struggle*. Berkeley, CA: University of California Press.

Platt, John. 1999. "Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods." *Advances in large margin classifiers* 10 (3): 61–74.

Popkin, Samuel L. 1979. *The Rational Peasant: The Political Economy of Rural Society in Vietnam.* Berkeley, CA: University of California Press.

Popkin, Samuel L. 1988. "Political entrepreneurs and peasant movements in Vietnam." *Rationality and revolution* pp. 9–62.

Regan, Patrick M, and Daniel Norton. 2005. "Greed, grievance, and mobilization in civil wars." *Journal of Conflict Resolution* 49 (3): 319–336.

Robinson, James A, and Ragnar Torvik. 2005. "White elephants." *Journal of Public Economics* 89 (2): 197–210.

Ross, Michael L. 2004. "What do we know about natural resources and civil war?" *Journal of peace research* 41 (3): 337–356.

Rossmo, D Kim. 1999. *Geographic profiling.* CRC press.

Saif, Hassan, Yulan He, Miriam Fernandez, and Harith Alani. 2014. Semantic patterns for sentiment analysis of Twitter. In *International Semantic Web Conference.* Springer pp. 324–340.

Sambanis, Nicholas. 2001. "Do ethnic and nonethnic civil wars have the same causes? A theoretical and empirical inquiry (part 1)." *Journal of Conflict Resolution* 45 (3): 259–282.

Sands, Melissa L. 2017. "Exposure to inequality affects support for redistribution." *Proceedings of the National Academy of Sciences* p. 201615010.

Schneider, Mark, and Paul Teske. 1992. "Toward a Theory of the Political Entrepreneur: Evidence From Local Government." *American Political Science Review* 86 (03): 737–747.

Schussman, Alan, and Sarah A Soule. 2005. "Process and protest: Accounting for individual protest participation." *Social forces* 84 (2): 1083–1108.

Smith, Douglas A. 1986. "The neighborhood context of police behavior." *Crime and justice* 8: 313–341.

Snyder, David, and Charles Tilly. 1972. "Hardship and collective violence in France, 1830 to 1960." *American Sociological Review* pp. 520–532.

Steinert-Threlkeld, Zachary C. 2017. "Spontaneous Collective Action: Peripheral Mobilization During the Arab Spring." *American Political Science Review* .

Steinert-Threlkeld, Zachary C., Delia Mocanu, Alessandro Vespignani, and James H. Fowler. 2015. "Online Social Networks and Offline Protest." *EPJ Data Science* 4 (19): 1–9.

Stevenson, M.D, Verity, and R. 2014. *Rgeoprofile : Geographic Profiling in R.* London, England: Queen Mary University of London. Version 1.2.
URL: *http://evolve.sbcs.qmul.ac.uk/lecomber/sample-page/geographic-profiling/*

Stewart, Frances. 2008. *Horizontal Inequalities and Conflict: An Introduction and Some Hypotheses.* Palgrave Macmillan chapter Horizontal Inequalities and Conflict: An Introduction and Some Hypotheses, pp. 3–24.

Tarrow, Sidney. 1994. *Power in movement: Social movements, collective action and mass politics.* New York, NY: Cambridge University Press.

Thomson, Henry. 2016. "Rural Grievances, Landholding Inequality, and Civil Conflict." *International Studies Quarterly* 60 (3): 511–519.

Tullock, Gordon. 1971. "The Paradox of Revolution." *Public Choice* 11: 89–99.

Tyler, Tom R, Jeffrey Fagan, and Amanda Geller. 2014. "Street stops and police legitimacy: Teachable moments in young urban men's legal socialization." *Journal of Empirical Legal Studies* 11 (4): 751–785.

Uggen, Christopher, and Jeff Manza. 2002. "Democratic contraction? Political consequences of felon disenfranchisement in the United States." *American Sociological Review* pp. 777–803.

Verity, Robert, Mark D Stevenson, D Kim Rossmo, Richard A Nichols, and Steven C Le Comber. 2014. "Spatial targeting of infectious disease control: identifying multiple, unknown sources." *Methods in Ecology and Evolution* 5 (7): 647–655.

Walls, David S. 2015. *Community organizing.* San Francisco, CA: John Wiley & Sons.

White, Ariel. 2015. "Misdemeanor Disenfranchisement? The demobilizing effects of brief jail spells on potential voters." Harvard Scholar.

White, Ariel R, Noah L Nathan, and Julie K Faller. 2015. "What do I need to vote? Bureaucratic discretion and discrimination by local election officials." *American Political Science Review* 109 (1): 129–142.

Wildeman, Christopher. 2014. "Parental incarceration, child homelessness, and the invisible consequences of mass imprisonment." *The Annals of the American Academy of Political and Social Science* 651 (1): 74–96.

Willer, Robb. 2009. "Groups reward individual sacrifice: The status solution to the collective action problem." *American Sociological Review* 74 (1): 23–43.

Williamson, Vanessa, Theda Skocpol, and John Coggin. 2011. "The Tea Party and the remaking of Republican conservatism." *Perspectives on Politics* 9 (1): 25–43.

Wood, Elisabeth Jean. 2003. *Insurgent collective action and civil war in El Salvador.* New York, NY: Cambridge University Press.

Zuern, Elke. 2011. *The Politics of Necessity: community organizing and democracy in South Africa.* Madison, WI: University of Wisconsin Press.

6 Appendix

6.1 Figures

6.1.1 ROC Plot

As Figure 9 shows, the false positive rates for our LR and SVC models are similar.

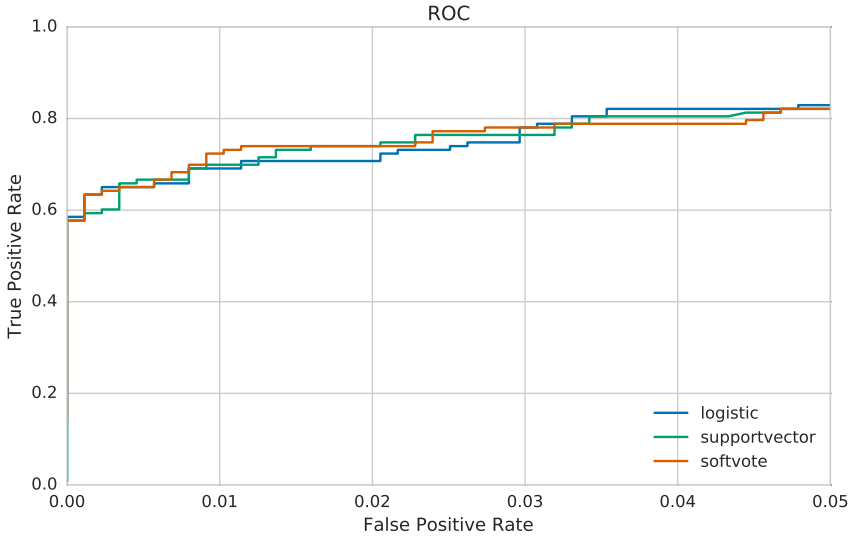

Figure 9: This plot shows the receiver operating characteristic (ROC) curve from the three estimated models. When using a soft vote model with a false positive rate of 0.01, our true positive rate is approximately 80%.

6.1.2 Feature Matrix

Figure 10 graphs the feature matrix from our hand-coded training set via t-distributed stochastic neighbor embedding (t-SNE). The t-SNE algorithm maps the high-dimensional feature matrix into two-dimensional space while trying to preserve the local distance structure for visualization. Although the axes are unintereptable, t-SNE plots are useful to visualize the amount of overlap among features. Appendix Figure 10 suggests that the features from protest-related and non-protest related twitter accounts are not immediately separable, motivating our use of a classifier algorithm.

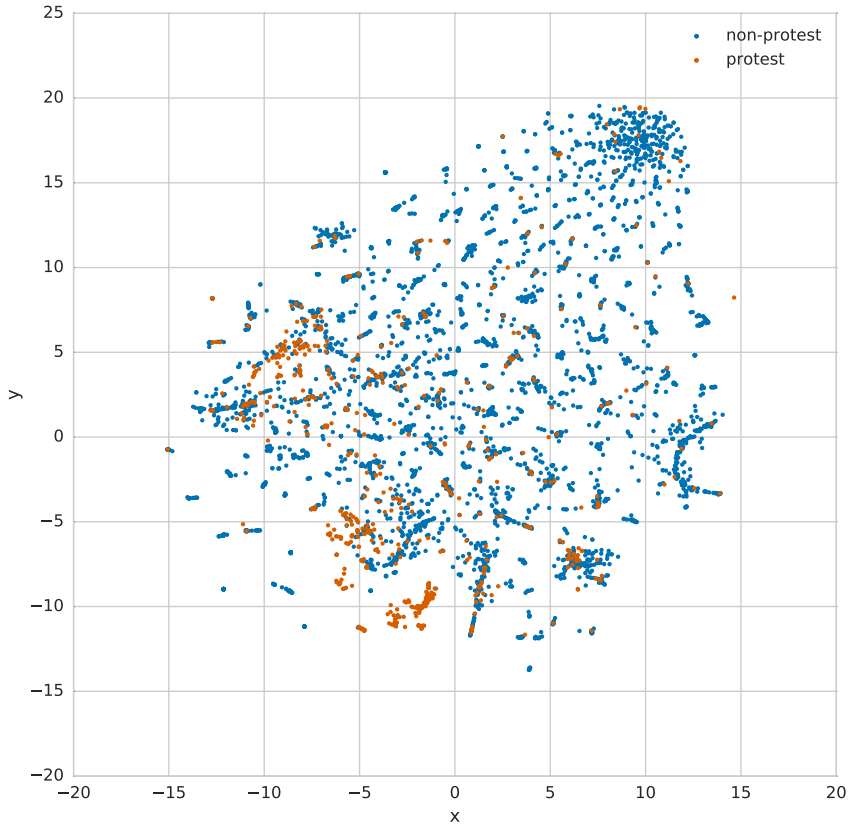

Figure 10: t-SNE plot that collapses the high-dimensional feature matrix into two dimensions. Features from protest-related accounts are denoted in orange, while non-protest related accounts are denoted in blue.

6.1.3 Density of Control Variables

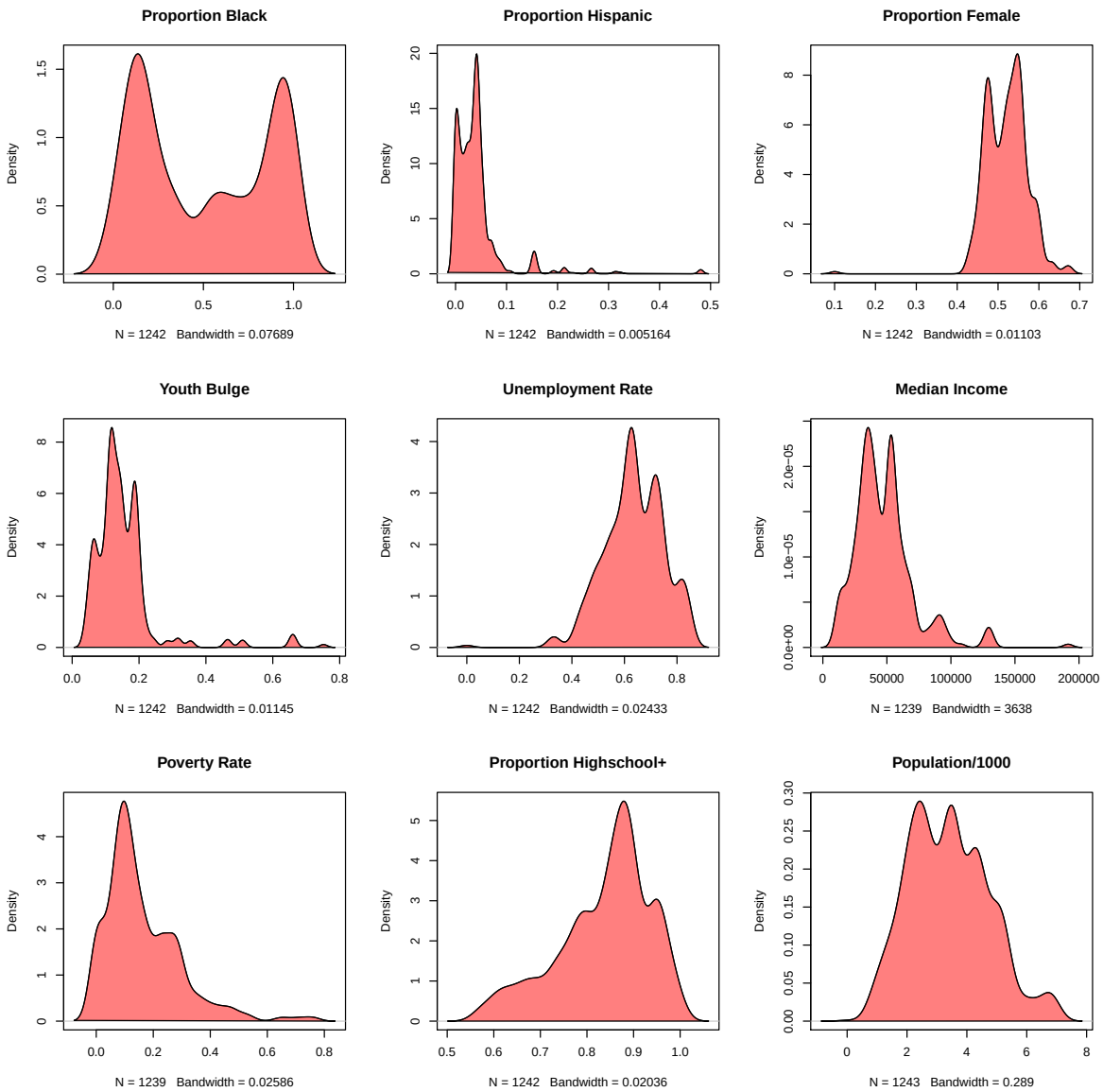

Figure 11: Density of the control variables in the DPM models.

6.2 Algoritms

6.2.1 Algorithm I

Algorithm 1: Social Activity Hub Estimation for Each User

Data: The set of n observed incidents $\mathbf{x}_i = (x_i^{(1)}, x_i^{(2)})$, $i = 1, ..., n$
if $n = 1$ **then**
  Assign $\mathbf{x}_1$ as $\text{SAH}^{(\text{sole observed})}$;
else
  MCMC algorithm implemented as the `RunMCMC()` function in Rgeoprofile 1.2 (Stevenson et al. 2014: summarized below in algorithm 2), based on the discussion in Verity et al. (2014));
  if *convergence fails;*
  then
    if $n > 25$ **then**
      remove incidents outside of bounding box created by the 1st and 99th percentiles of $x^{(1)}$ and $x^{(2)}$;
    Assign $\mathbf{x}_i$ closest in Euclidean distance to the spatial mean of $\mathbf{x}$ as $\text{SAH}^{(\text{spatial mean})}$;
  else
    Take 2000 posterior draws; thin by keeping the first of every 20;
    begin local minima model:
      combine all 100 posterior draws;
      a) calculate hit score surface;
      b) find local minima j on surface within 0.05 degree decimal radius, $j = 1, ..., \infty$;
      c) **foreach** *local minimum* j **do**
        Assign $\mathbf{x}_i$ closest in Euclidean distance to local minimum as $\text{SAH}_j^{(\text{local minima})}$
      end
      foreach *posterior draw* **do**
        a to c;
      end
    end
    begin cluster mean model:
      combine all 100 posterior draws;
      d) **foreach** *cluster* j *of* $\mathbf{x}$ **do**
        Assign $\mathbf{x}_i$ closest in Euclidean distance to estimated source of cluster as $\text{SAH}_j^{(\text{cluster mean})}$
      end
      foreach *posterior draw* **do**
        d;
      end
    end
  end
end
Result: **SAH**=($\text{SAH}^{(\text{sole observed})}$, $\text{SAH}^{(\text{spatial mean})}$, $\text{SAH}^{(\text{local minima})}$, $\text{SAH}^{(\text{cluster mean})}$)

6.2.2 Algorithm II

Algorithm 2: `RunMCMC()` from Rgeoprofile 1.2 (Stevenson et al. 2014)

Data: The set of n observed incidents $\mathbf{x}_i = (x_i^{(1)}, x_i^{(2)})$, $i = 1, ..., n$
Initialize by setting initial values and computing relevant priors;
Define sampling steps:
  a) draw and update $\mathbf{z}_{c_i}$ based on most updated c_i;
  b) draw and update c_i based on most updated $\mathbf{z}$;
begin Burn-in
  repeat
    for *i in* 1 *to n* **do** a-b;
  until *convergence*;
end
begin Posterior draws
  foreach *posterior draw* **do**
    for *i in* 1 *to n* **do** a-b;
  end
end
Result:
  1. For each $\mathbf{x}_{1,...,n}$, its corresponding cluster c_i
  2. For each unique cluster c_j, its source location $\mathbf{z}_j$

6.3 Alternate Estimation of Social Activity Hubs

In the main body of the paper, we estimate each user's social context using our novel social activity hubs (SAHs) approach. This has the benefit of allowing users' context to range across multiple places within Baltimore. However, it is possible that our results are an artifact of some as-yet unknown bias in this estimation procedure. In Table **??**, we explore this possibility by assigning each user's social activity to their the modal tweet location, i.e. the place from which they tweet most often. Our results in this table are almost entirely consistent with our model estimated using our SAH measure.

Table A1: Models of Protest Participation; Home Locations by Modal Tweet

	Model III		Model IV	
	Estimate	S.E.	Estimate	S.E.
Exposure to Policing†	0.10*	0.04	0.11*	0.05
Proportion Black	−0.33	0.27	−0.32	0.27
Proportion Hispanic	−2.71	1.41	−2.71	1.41
Proportion Female	0.61	1.51	0.62	1.51
Youth Bulge	−0.30	0.75	−0.31	0.75
Median Income	1.00	0.82	1.01	0.82
Poverty Rate	0.00	0.00	0.00	0.00
Unemployment Rate	−0.33	0.71	−0.32	0.71
Proportion HS Graduate	−1.40	1.11	−1.43	1.11
Per 1000 Population	0.03	0.21	0.03	0.21
Per 1000 Population^2	−0.02	0.03	−0.02	0.03
Intercept	0.16	1.37	0.17	1.37
N	1,501		1,501	
AIC	1863.5		1863.3	

* $p < 0.05$
† Exposure to policing is per 100 arrests in Model III and per 100 visits in Model IV.

6.4 Other Non-Grievance Inducing Arrests

In our mechanism section, we argue that arrests where officers can show discretion generate grievances, while arrests supported by the community do not. In the main body, we show that arrests for murder or for crimes that carry a maximum penalty of life do not affect participation in collective action. In this section, we consider whether our evidence behind our mechanism depends upon our choice of the type of crime we consider "community-supported." Specifically, we also explore whether arrests for "personal" crimes such as assault, abuse, and kidnapping affect participation. We also explore whether crimes where a weapon was used affects participation. The results, summarized in Tab A2, do not suggest our choice of crime affects our results, as arrests are not significant in any model.

Table A2: Models of Protest Participation Using Social Activity Hub: Drug Arrests

	"Personal" Crimes				Armed			
	Model I		Model II		Model I		Model II	
	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Exposure to Policing†	0.14	0.10	0.17	0.11	0.48	0.38	0.57	0.44
Proportion Black	0.07	0.30	0.07	0.30	0.08	0.31	0.07	0.31
Proportion Hispanic	-3.07	1.75	-3.07	1.75	-3.06	1.75	-3.06	1.75
Proportion Female	-0.55	1.59	-0.51	1.59	-0.84	1.55	-0.83	1.55
Youth Bulge	0.74	0.80	0.73	0.80	0.69	0.82	0.67	0.82
Unemployment Rate	-1.15	0.90	-1.14	0.90	-1.24	0.91	-1.25	0.91
Median Income	-0.00	0.00	-0.00	0.00	-0.00	0.00	-0.00	0.00
Poverty Rate	0.09	0.73	0.09	0.73	0.11	0.73	0.10	0.73
Proportion HS Graduate	-0.24	1.23	-0.25	1.23	-0.09	1.21	-0.09	1.21
Per 1000 Population	0.31	0.25	0.31	0.25	0.32	0.25	0.32	0.25
Per 1000 Population^2	-0.06	0.03	-0.06	0.03	-0.06	0.03	-0.06	0.03
Intercept	0.09	1.48	0.07	1.48	0.21	1.49	0.21	1.49

† Exposure to policing is per 100 arrests in Model I and per 100 visits in Model II.

6.5 Variable selection

In this section, we present the results from a least absolute shrinkage and selection operator (Lasso) regression. Lasso regressions estimate the effect of reducing variables to zero on the models prediction accuracy. While the interpretation of the coefficients is the same as standard logistic models, Lasso regression identify which covariates do not reduce prediction accuracy when omitted from the model. After crossvaligdation, we find that omitting both measures of arrests, proportion Hispanic, poverty rate, and population^2 significantly reduce prediction error. Other variables can be omitted. These results are presented in Table **??**.

Table A3: Lasso Regressions of Protest Participation Using Social Activity Hub

	Model I	Model II
	Coef.	Coef
Exposure to Policing†	0.06	0.05
Proportion Black		
Proportion Hispanic	−2.12	−2.12
Proportion Female		
Youth Bulge		
Median Income		
Poverty Rate	0.38	0.38
Unemployment Rate		
Proportion HS Graduate		
Per 1000 Population		
Per 1000 Population^2	−0.01	0.01
Intercept	−0.64	−0.64

† Exposure to policing is per 100 arrests in Model I and per 100 visits in Model II.

Note that standard errors are not available from Lasso estimates.

6.6 Influential Observations

One concern is that our results are driven by a few influential observations that bias our estimates. To assess whether this is the case, we perform two regression diagnostic tests. In the first, we estimate the DF Betas (D_{ij}) for our measure of exposure to policing. A DFBeta is defined as $D_{ij} = \beta_j - \beta_{j(-i)}$ for $i = 1, \ldots, n$ and $j = 0, 1, \ldots, k$ where β_i are for all observations and $\beta_{j(-i)}$ are those with the ith observation removed. Appendix Figure 12 shows that almost all observations lie along a single line, with no observations exceeding the critical value.

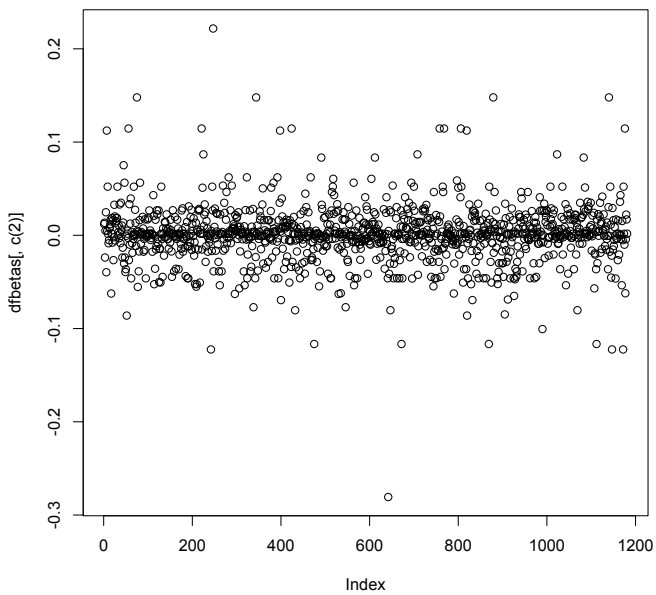

Figure 12: Figure of the DF Betas estimated for our first model.

To confirm this observation, we also estimate Cook's distance, which estimates an F-test for the hypothesis that $\beta_j = \beta_{j(i-)}$ for $j = 0, 1, \ldots, k$. Although there is no significance test for Cook's distance, the rule of thumb cutoff is $D_i > \frac{4}{n-k-1}$. The results from this estimation are presented in Appendix Figure 13. While three observations appear to exceed the critical value, omitting them from our analysis does not change any results.

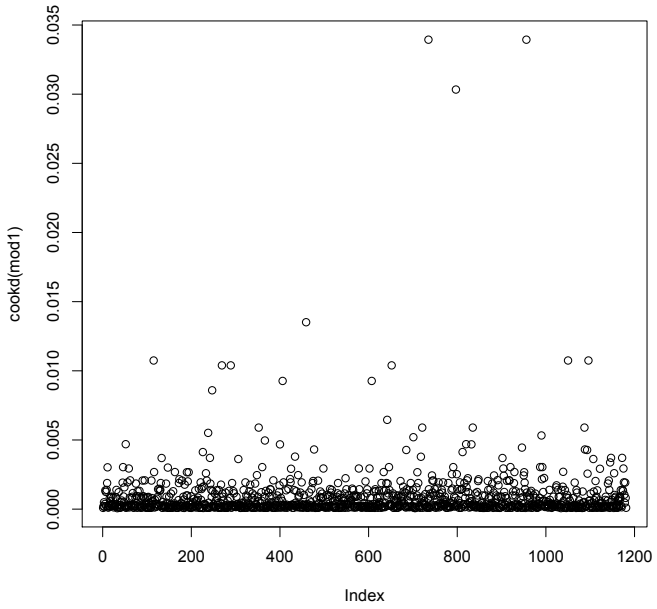

Figure 13: Figure of Cook's distance estimated for our first model.