Organization, Repression, and the Violent Escalation and De-Escalation of Nonviolent Protest

Vilde Bergstad Larsen

Master thesis in Political Science

Institute of Political Science

Department of Social Sciences

32 948 words

Spring 2020

2020

Organization, Repression and the Violent Escalation and De-Escalation of Nonviolent Protest

Word count: 32 948

http://www.duo.uio.no

Abstract

Why do some protests turn violent while others do not? The violent escalation of demonstrations is subject to massive media coverage, yet little empircal research on the causes of this has been done. This thesis offers a spatially and temporally disaggregated research design that analyzes the dynamics of protest events in Africa and Asia from 2010-2018. The results from various specifications of the OLS regression model show that government repression of nonviolent protest preludes violent escalation. Repression of violent protest has the opposite effect, instigating violent protesters to switch to nonviolent means. Furthermore, I find that the association between repression and violent escalation is conditional upon whether nonviolent protests are organized or not. Even in the face of repression, organized nonviolent protests are more effective at maintaining nonviolent discipline than their disorganized counterparts.

Acknowledgements

"Until we are all free, we are none of us free."

Emma Lazarus

This project started out with the Hong Kong protests and the worldwide Fridays for Future climate strikes. It ended with a wave of demonstrations in response to racial injustice and police brutality. Throughout this year, I have been sitting comfortably behind the computer, breaking very real grievances, unjust and oppression down into data points – at a safe distance. I feel extremely privileged, yet uncomfortably numb. This thesis is dedicated to protesters who remain dedicated to the nonviolent cause, sacrificing their personal safety for causes that have the potential to make the world a better place; politicians who listen to the voices of the people; and police forces who don't respond to nonviolence with violence.

First, I want to thank my supervisors, Professor Tore Wig (UiO) and Senior Researcher Siri Aas Rustad (PRIO), for pushing me across the finishing line. Thank you for your valuable input, guidance, and for believing in this project. 10/10 would recommend.

Second, I owe a huge thanks to everyone at the Peace Research Institute of Oslo (PRIO) for providing me with a writing spot in an open and inspiring academic environment, for all the help with setting up a home office when the pandemic hit, and for the endless coffee. A special thanks to Senior Researcher Marianne Dahl and Tora Sagård for taking their time to read and discuss my thesis with me.

Third, I am extremely grateful for all my dear friends; thank you for reminding me that there is so much more to life than thesis writing and for making even the most stressful times good times. A special thanks to vbj Mona, Ingrid and Ingeborg for proof-reading both my thesis and my Tinder messages.

Finally, I want to thank my family. To my siblings, Eivin, Siril and Kine – thank you for setting fantastic examples, both in pursuing higher education (although our academic interests greatly differ) and in life, as well as for giving me six wonderful nieces and nephews. To my parents, thank you for teaching me how to learn, for always supporting me, and for allowing me to explore the world. I (literally) wouldn't be here without you.

Any mistakes are solely my own.

R scripts are available on GitHub.com/ma_vbl.

Oslo, June 22, 2020

Vilde Bergstad Larsen

Contents

List of Tables

List of Figures

1 Introduction

On Wednesday, October 14, 2015, students at the University of the Witswatersrand, called Wits University, started what turned into a nation-wide protest movement against academic and financial exclusion from higher education. Wits University is one of the largest public universities in South Africa, located just outside of Johannesburg. Earlier that week in October 2015, the university sector had announced that tuition fees were to increase by over ten percent. That was the spark that caused the prairie fire of student-led protests across South Africa in 2015 and 2016.

The angry students at Wits gathered behind the Twitter hashtag #WitsFeesMustFall. However, the fee increases were not exclusive to this one university. Within a matter of days, the fee protests spread from Johannesburg to several other universities across the country (Ngcobo 2015). Now, the students started to rally behind the slogan #FeesMustFall (FMF). By the following Monday, academic activities were suspended at several universities after increased tension. During the first month of fee protests, the events were described as organized and peaceful. Gradually, the protests fizzled out at some universities, while others "erupted into mayhem" (Poggi 2015).

The aim of this thesis is to explain the violent escalation and de-escalation that occurs during protest. Why did the protests at some universities fizzle out, while increasing violent protest occurred elsewhere? While violent escalation of protest is a common phenomenon that often receives much attention in the media, the mechanisms that lead to this are understudied in the literature. To answer this question, I focus on two variables. The first variable, government repression, concerns the response of the government when met with dissent. The second variable, level of organization, describes a central characteristic of the protest activity.

Unsurprisingly, the #FMF protests gained increased media attention as they escalated in a violent direction, with reports of property damage, barrication of buildings, and clashes with security forces. In 2018, three years after the initial protests, a government report estimated that the total property damage caused by the #FMF protests at higher education institutions amounted to nearly 800 million Rand, or approximately 42 million US Dollars (Dentlinger 2018). Indicative of the geographical spread, the same report revealed that only two out of the country's 13 universities did not report damages to property due to the protests.

The South African government were quick to announce that tuition fees would not increase in 2016 as planned (Pearson, Karimi, and McKenzie 2015). At some campuses, academic activities resumed and calm returned, while protests continued elsewhere. The level of violence in the protests also varied within and between universities. One year later, in September 2016, however, protests flared up again at campuses across the country (Roberts 2016). The underlying causes of the new of wave protests remained the same, sparked by an announcement made by the government of an eight percent fee increase the coming academic year. The 2016 protest wave has been described as qualitatively different from the protests in 2015 along many dimensions, including the measures used to repress the protests and the severity of property destruction (Ndlovu 2017, 136). Again, there were considerable differences between universities, both in overall protest level and in the level of violent protest activity.

Figure 1 shows how protests spread across South Africa in the first three weeks of the 2015 wave, from October 12 to November 1. In the first week, the cluster of protests in the south-west are located in and around Cape Town, while the north-eastern protest cluster is around Johannesburg, the location of the Wits University. Protests spread quickly from these two hotspots to other parts of the country, as evident from plots of the second and third week. The size of the symbols represent the number of unique protest events that occurred during one week. From these simple maps, the subnational variation is striking. Protest at different locations have unique situational dynamics, which is the cornerstone of what I explore in this thesis.

The #FMF protests illustrate the two key themes of this thesis. First, there was considerable interaction between the protesters and police and security forces. Twitter videos document the chaotic scenes when security guards were called in to contain the demonstrations (eNCA 2015). Throughout the protests, police forces were reported to have used violent means, such as tear gas, stun grenades, and rubber bullets, in attempts to disperse the crowds at several universities (Eyewitness News 2015). During the 2016 wave, the police brutality was compared to tactics used by the apartheid government (Ndlovu 2017, 136). One of the questions that remains debated in the literature is when the use of such repressive measures is effective, in terms of quelling or deterring dissent, and when it is not. In this thesis, one of the factors that I explore is this relative effectiveness of government repression of demonstrations, that is, whether it is linked to more or less use of violence by the protesters. Specifically, I look at whether the violent escalation of protest is influenced by preceding repression.

Second, the protests started as spontaneous reactions to the tuition fee increase[1]. Apart from the Twitter hashtag, there is little evidence of a formal, organizational structure underpinning and coordinating the demonstrations. A common assumption made in the literature is that clear leadership is important to maintain nonviolent discipline within the protest (Pinckney 2016). Relatedly, government repression is found to be less effective in quelling organized dissent (Chenoweth, Perkoski, and Kang 2017). However, little research has been made on more spontaneous and relatively disorganized protests. The second set of factors that I study in this thesis therefore relates to the level of organization "on the ground". Specifically, I ask whether the risk of increasing violence in an area is higher where fewer protests are organized.

[1] There were a multitude of underlying causes to the protests, but the announcement of fee increases was the precipitating factor (Ndlovu 2017). At its core, the protests symbolized the structural racism, embedded in academic and financial exclusion, still evident in post-Apartheid South Africa. Only a few months prior, similar mass protests had broken out at the University of Cape Town (UCT) calling for, and ultimately succeeding in, the removal of a statue of the British imperialist Cecil J. Rhodes on campus (Ndlovu 2017).

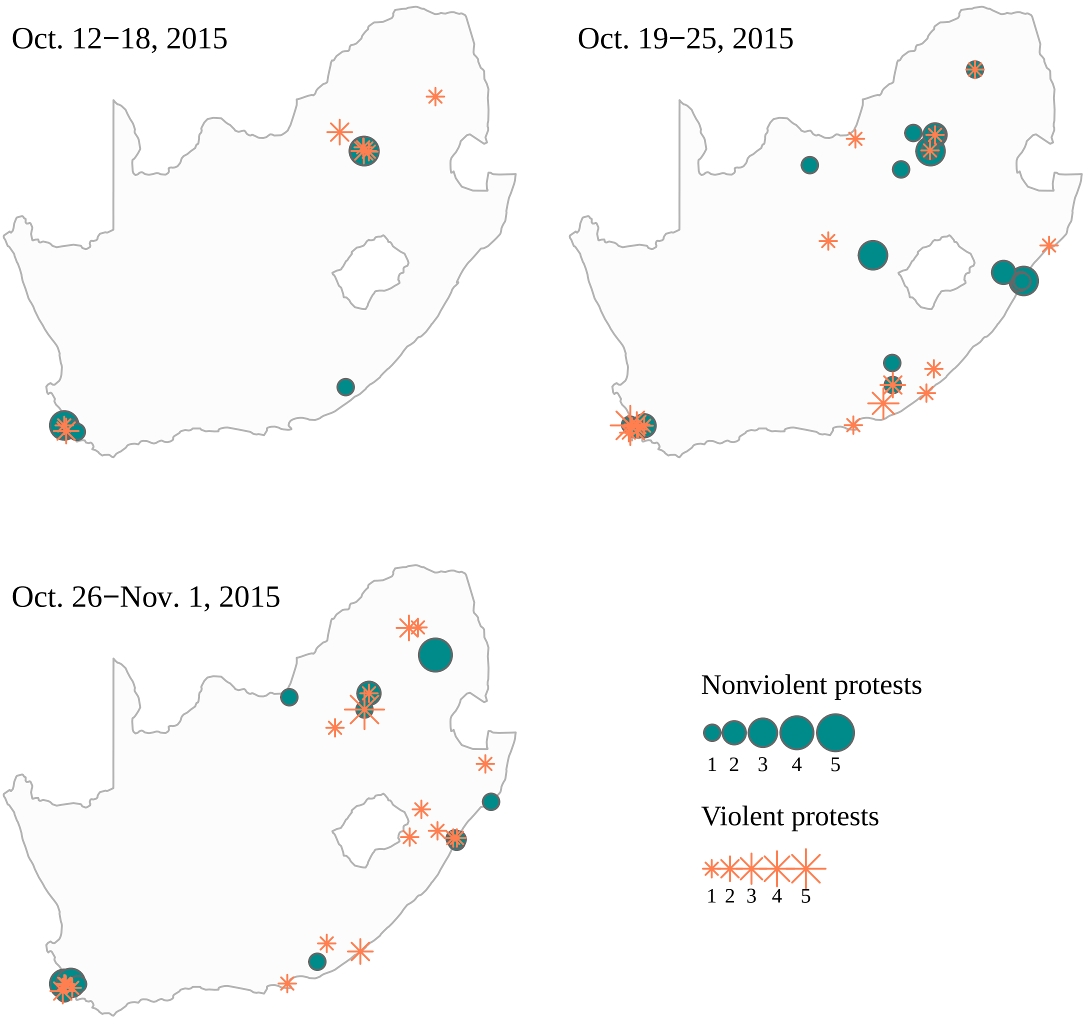

Figure 1: Protest events in South Africa, October 12-November 1, 2015 (ACLED)

1.1 Research question and delimitations

In this thesis, I pose the following research question:

Why do some protests escalate violently while others remain nonviolent?

To answer this question, I am specifically looking at variation in government repression of protests and the level of organization. These variables account for the dynamic interaction between protesters and governments. Repressive behavior be considered a central characteristic of the government when facing nonviolent or violent dissent. Whether protests are organized or not, in the sense that an organizational body is behind the demonstration, taps into a core characteristic of the protest. I present a Rational Actor (RA) model of individual protest behavior. The model shows how both government repression and protest organization influence the protester's calculation of the relative costs and benefits of resorting to violence in comparison to remaining nonviolent.

Large bodies of research have studied either the onset of nonviolent protest or the onset of intrastate political violence. The transition from nonviolence to violence during the course of protests is a comparatively understudied area. This is despite the fact that such escalation is relatively common and is subject to massive media coverage.

Answering the research question of this thesis requires an understanding of the sequence of dynamic interaction that happens in between onset and outcome. Hence, this thesis is placed in the methodological field of large-N protest event analysis (Rucht 2007). I use weekly event data on protest from the Armed Conflict Location and Event Data Project (ACLED) to answer the question. In around 17 percent of the weekly observations in my dataset ($N = 20232$) there was an *increase* in the level of protest violence compared to the previous week. Vice versa, *decreasing* levels of violence are observed in 10.1 observations.

I focus on the type of protests that take the form of *demonstrations*, as opposed to individual statements of dissent, where people take to the streets to publicly express their claims. These are primarily symbolic actions that seek to communicate an opinion in order to influence an audience (Sharp 1973, 117–18). The terms protest and demonstrations are used interchangeably throughout this thesis.

Building upon the literature on other methods of violent and nonviolent resistance, I view protest as a sub-class of these more general categories. It can be contended that a full analysis of the shift from nonviolence to violence should include other methods of nonviolent resistance, such as sit-ins and strikes (see Sharp (1973) for a comprehensive description of the methods of nonviolent action). Similarly, the repertoire of violent action could also have included i.e. terrorist attacks or guerrilla warfare. I have chosen against this and restrict the analysis to protest for both substantial and pragmatic reasons. On a practical note, it is a choice based on data availability, as well as the limited scope of this thesis. In addition, the 2010s saw a remarkable incrase in the number of protest events across the globe. Understanding the dynamics of these events is important in itself, although I acknowledge that protest may be accompanied by other methods of civil resistance.

1.2 Structure of thesis and key findings

In this section, I present the structure of the remainder of my thesis. In the remainder of this chapter, I present the contributions this thesis makes to the literature and the policy relevance of the research question. In Chapter 2, I review the literature on protest and intra-state political conflict, nonviolent and violent, as well as the persistent debate of the effects of repression. Furthermore, I demonstrate how unorganized demonstrations have been largely overlooked in the literature. I identify the knowledge gaps in the literature, and explain how this thesis furthers the relevant research agenda on tactical shifts during protest. In Chapter 3, I present the Rational Actor (RA) model that is the theoretical framework of this thesis. I explain how and why government repression and organization influence the choice between nonviolent and violent protest through altering the individual protester's cost-benefit calculation. I also present the hypotheses that are tested in the empirical analysis. Following the theory chapter, I present the data source, how the dataset was constructed, and operationalizations of key variables in Chapter 4. In this chapter, I also present the research design which is based on estimating various specifications of the linear regression model in order to assess the robustness of the results. In Chapter 5, I turn to the empirical analysis of the data. I demonstrate that mild repression of nonviolent protest is associated with an increase in the proportion of violent protest in the subsequent week. Repression of violent protest and high levels of organization is shown to be consistently associated with decreasing levels of violent protest in the following week. The results are found to be robust across different models. Chapter 5 also includes a discussion of the internal and external validity of the regression models, and of whether the results can be interpreted in causal terms. I conclude the thesis in Chapter 6, where I summarize my findings and point to avenues for further research.

1.3 Relevance and contribution: Repression and organization

This thesis is important from both an academic and a policy perspective.

Academically, I situate the thesis within the growing body of research that directly addresses tactical shifts that occur *during* a protest. For a long time, the study of protest and civil dissent was separated by a nonviolent and violent axis. Recently, this separation has been put under increasing scrutiny (Asal et al. 2013). This development is coupled with a recognition that tactical choices by dissident actors are indeed strategic choices; purposively chosen means to an end.

In Chapter 4, I use a Rational Actor framework and develop a theoretical model that explains how repression and level of organization influences increasing use of violent tactics. In much of the literature on protest, the campaign, defined as "a series of observable, continual tactics in pursuit of a political objective" (Chenoweth and Stephan 2011, 14), is used as the central unit of analysis. Campaigns are usually named, have discernible beginning and end points, and have a clear leadership structure (Chenoweth and Stephan 2011). I offer an alternative approach to the study of protest behavior, inspired by Cunningham and Beaulieu (2010). Instead of asking the question: "Was the movement nonviolent or violent?" or "Did the movement change its primary tactic?", I offer a more nuanced perspective by analyzing the relative use of violent tactics at a specific point in time and space.

Furthermore, when analyses are situated on the campaign- or organization-level, a potential selection bias is left largely unaddressed. The definition of the campaign entails that there is a certain level of organization to the protest. Thus, the spontaneous and disorganized sparks that may eventually escalate to large protests or revolutions is, by definition, excluded from much of the scholarly work on protest. I offer a disaggregated analysis where events at specific points in time and space, rather than the organization or campaign, take center stage. This allows me to also look at relatively disorganized protests. Taking this one step further, I explore whether there exists a link between the level of organization "on the ground", in terms of whether protests are organized by a named group or not, and the level of violence employed by the protesters. This is an aspect largely left unexplored in the literature due to the focus on campaigns.

To exemplify, during the course of the #FMF protests, grievances, demands, and, importantly, methods varied within the protest and from university to university (Ndlovu 2017). While dramatic incidents and violent clashes received much attention in the media, most of the student protesters were, and remained, nonviolent (Ndlovu 2017). I contend that treating protests such as the #FMF as unitary nation-level protest campaigns is problematic precisely because of the considerable situation-specific differences. In this thesis, I provide insight into the situational circumstances in which protests occur by offering a dynamic theoretical model coupled with a spatially and temporally disaggregated research design. This explicitly recognizes that the level of violent protest can vary within a country. Furthermore, as protests are generally short-lived phenomena, tactical changes are modeled on a week-by-week basis and the core explanatory factors are measured accordingly.

Turning to the policy-relevance of this thesis, the 2010s has been called the decade of protests (Younge 2019). Nonviolent and violent protests have mobilized citizens all around the world, shaping domestic and international politics alike. This thesis therefore contributes to the understanding of how the dynamics of protests are influenced both by the government and the protesters themselves.

The literature on outcomes of violent conflict has shown that the negative effects of phenomena such as insurgency, revolutions and civil war, tend to last for a long time, affecting both the socioeconomic and political life of society (Davenport et al. 2019). While violent protest does not necessarily produce the same devastation to society as civil war in itself, the 2011 Arab Spring is evidence that the line between nonviolent and violent protest, revolution and civil war can be fine. Furthermore, violent escalation of protest may impose significant costs on society resulting from property damage as in the aftermath of the #FMF protests. History has thus provided sufficient evidence that violence should be avoided - not only from a moral perspective, but also from a socioeconomic perspective.

The results from this thesis show that government repression of nonviolent protest is associated with a significant increase in the proportion of violent protest in the following week. In other words, violent protest escalation is largely influenced by the actions of the government. If the association between repression and violent escalation is generalizable, this is a very concrete example of why governments should carefully consider their responses to nonviolent protesters. Furthermore, violent protests are by definition more disruptive to society and, as shown by Chenoweth and Stephan (2011), movements that are able to remain nonviolent have a higher success rate than their violent counterparts. If I am correct in that a certain level of organization is important for maintaining this nonviolent discipline - also in the face of repression, this thesis also provides lessons for activists as violent escalation reduces the public support of dissenting groups, in turn decreasing their chances of success (Simpson, Willer, and Feinberg 2018),

In this chapter, I have presented the research question and the two explanatory variables that I focus on in my thesis: repression and organization. In addition, I have answered the “So what?” question that is imperative to ask in any research project, where I argued that this study furthers the research agenda on nonviolent and violent protest and is also important from a policy perspective. I now turn to the literature review where I go more in detail on the knowledge gaps.

2 Nonviolent and violent protest: A review of the literature

In this chapter, I begin with a brief discussion of the distinction between nonviolence and violence. Thereafter, I discuss protest as a form of contentious politics, whether nonviolent or violent. It is natural to place the specific phenomenon of protest within the larger bodies of research on non-institutional political conflict, particularly regarding violent protest which has tended to be studied separate from nonviolent protest. Episodes of contention vary considerably in duration, but all protests have a start and an end point. These two distinct stages are what has received most attention in the literature. Comparatively less research has been made on shifts from nonviolence to violence, or vice versa, during the course of a protest.

I continue the literature review with a discussion of the shortcomings of traditional structural models of political violence onset. While valuable in identifying patterns of the onset of internal political conflicts, I show that these studies do not address the move from nonviolence to violence nor the dynamic interaction between protesters and governments.

Increasingly, however, scholars are beginning to empirically analyze the dynamics of contention. I situate my thesis within this expanding sub-field of protest research. In the final section, I place particular emphasis on the diverging empirical findings on the link between government repression and violent escalation of protest. I find that existing literature almost exclusively analyzes organized protest movements and campaigns, leaving little attention to relatively disorganized protest events. I therefore conclude the literature review by spelling out the knowledge gaps that this thesis addresses: the links between repression, level of organization, and the violent escalation and de-escalation of nonviolent protest.

2.1 Defining violence and nonviolence

The research question of this thesis is what leads to changes in the level of violence in initially nonviolent protest. It is therefore important to clearly define what I mean with violence and nonviolence.

I follow the conventional definition of violence as the intended physical damage to persons or property (Bond et al. 1997; Pinckney 2016). In this definition, there is no minimal threshold of fatalities or injuries required for a protest event to be characterized as violent, nor need weapons be used. While some contend that property damage should not be considered violent resistance (Sharp 1973, 608), Pinckney (2016) argues that "even actions with minimal actual harm may be perceived as harmful and threatening" (p. 16).

I have chosen to define nonviolence in negative terms, as the absence of intentional physical harm toward persons or property (Bond 1988). This provides a distinction between nonviolence and violence that is empirically grounded. In his seminal work, Sharp (1973) identifies 198 methods of nonviolent action. The overview includes well-known types of civil resistance, such as sit-ins, strikes, boycotts, and protest, which is the focus of this thesis. The absence of harm does therefore not mean passive inaction. Rather, nonviolent action is conventionally defined as the active *and* nonviolent "collective pursuit of social or political objectives" (Schock 2003, 705).

Historically, "nonviolence" has often been attributed with a normative dimension (Bond 1988). In the literature, the term has been used to refer to an ideologically, religiously or ethically founded belief system that principally rejects the use of violence of any kind (Chenoweth, Perkoski, and Kang 2017).

Although I do not reject the moral dimension of nonviolence, I define the use of nonviolence (or violence) as a strategic choice. In the literature, nonviolent and violent resistance have often been viewed and analyzed as two separate concepts, despite the fact that the phenomena under study are largely the same (Schock 2015). One of the more recent developments in the literature is a step away from the nonviolence-violence dichotomy (Asal et al. 2013; Cunningham 2013; Cunningham, Dahl, and Frugé 2017). Asal et al. (2013) argue that different methods of resistance, including nonviolent and violent protest, should be conceptualized as an à la carte menu of which groups can choose, and where the choice is susceptible to change during the course of events.

Cunningham (2013) finds that the many of the same group-level factors have a positive effect on the likelihood that groups use nonviolence or violence. In other words, nonviolent and violent groups may take organized action based on the same causal mechanisms. It is therefore possible that the choice between nonviolence and violence is determined by the situation and environment in which the group finds itself. This finding is a strong argument to analyze the choices of strategy within the same framework, which is what I do in this thesis.

Defining movements simply by the absence or presence of violence is, however, not without caveats. Describing groups as an ideal type, fully nonviolent or fully violent, does more often than not oversimplify reality (Cunningham and Beaulieu 2010). Within nonviolent movements, there are often violent subgroups or more radical flanks (Chenoweth, Perkoski, and Kang 2017). Nonviolent and violent protest are therefore often used in conjunction with one another at the same point in time and space. To account for the use of mixed tactics, or the relative use of violence, I do not analyze the aggregate behavior of large organized and coordinated movements or campaigns. Instead, I focus on the situational and relational factors that influence the risk of shifts in the primary tactic of dissent. This can entail moves toward both increasing and decreasing use of violence.

2.2 Protest as contentious politics

Taking to the streets is clearly not the only way to express dissent. The conventional channels of political action, such as voting, are the regular and institutionalized ways through which citizens express their political views (Bond et al. 1997; Schock 2013). Importantly, prior to e.g. an election, formal, written rules specify how the outcome of the conflict of interest should be determined (Bond et al. 1997).

Protests take place outside these routine channels of expressing political opinion. They are therefore not controlled by the authorities the same way an election is. Because these channels are pre-established and thereby have a lower threshold for mobilization, it is considerable less likely that participation will prompt a repressive response from the state, even in cases of conflict of interest, thanks to the formalized rules of the game (Cunningham 2013). Hence, participation in demonstrations is regarded a costlier form of political behavior (Dahlum and Wig 2019).

To overcome this cost barrier, people need to be sufficiently motivated and also have the opportunity to protest. On the individual level, several factors are found to influence the likelihood of participation in nonviolent protest. First of all, the typical protester is young and educated (Dahlum and Wig 2017; Stockemer 2014). Politically active individuals that vote and are members of civil society organizations are also found to be more likely to participate in unconventional political action (Stockemer 2014). Perhaps unsurprisingly is also personal dissatisfaction with the government found to increase the likelihood of protesting (Stockemer 2014). Nonviolent mobilization is also found to occur in political environments where these routine channels of political participation have proven insufficient (Harris and Hern 2018).

From this, it follows that protest is a form of *contentious* politics, in the sense that it involves episodic (i.e. non-routine), public and collective conflicts of interest between claim-makers and their objects, in which the government is a stakeholder (McAdam, Tarrow, and Tilly 2011; Tilly 2003). In the study of contentious politics, the focus is on the interactions and dynamics between actors involved in contentious episodes (McAdam, Tarrow, and Tilly 2011; Tarrow 2015). This approach thus bridges research on different non-routine political activities, hereunder nonviolent and violent protest, but also terrorism, revolutions and civil wars, and contrast these with routine politics (McAdam, Tarrow, and Tilly 2011; Porta 2012). I follow Dahlum and Wig (2019) and conceptualize demonstrations in line with Tarrow's defintion of contentious politics: "coordinated, collective claims on authorities, made through public performances" (Dahlum and Wig 2019, 5).

This defintion of demonstrations entails that there is some minimum level of coordination among the protesters. However, unlike much of the contemporary literature on nonviolent and violent protest, I do not confine the analysis to so-called “protest campaigns”. These are defined as “a series of observable, continuous tactics” that are employed by a named non-state actor with a distinguishable leadership, toward a state actor (Chenoweth and Stephan 2010, 250). Despite being common, relatively unorganized protests that do not fall into this definition is an understudied area within the protest literature.

According to Bond et al. (1997), the contentiousness of an action depends on its disruptiveness, i.e. the extent to which the conflict of interest occurs outside of the “routine conflict resolution procedures of a political system” and thus leads to uncertainty (p. 556). From Max Weber’s classical definition of the state as a political entity that holds a legitimate monopoly over the means of coercion within a territory, it follows that contentious politics always involves a relation to the state, or government.

The state’s political power can be defined as having the available means to effectively “achieve or prevent the implementation of the wishes of the power-holder” (Sharp 1973, 7). Wherever people take to the streets to voice their discontent, and this action is sustained, it signals a withdrawal of the support, obedience, and cooperation upon which the political power of the state depends (Sharp 1973). By doing so, the protesters increase the government’s costs of preserving the status quo. A protest is therefore not only contentious, it is also *coercive*, which is defined by Bond et al. (1997) as “the extent to which an action threatens or imposes negative social, economic, political, or physical sanctions for noncompliance” (p. 557), which may or may not include the use of violence.

To summarize, demonstrations are examples of collective, rather than individual, contentious and coercive political action. The protesters are united by a specific grievance or demand, and taking to the streets is defined as a strategic activity with a social or political objective. The intentions of the protest acion is to evoke awareness and publicity, rally support for their claim, and mobilize additional supporters, as a means to produce the desired change (Sharp 1973).

2.3 Mobilizing for success

Even though the #FMF protests began nonviolently, protest violence quickly escalated at some campuses. For a long time, the dominant view within political science was that the threatened or actual use of violence is the most effective way for non-state actors to attain policy goals (Abrahms 2006; Chenoweth and Stephan 2011). Indeed, Addison (2002) describes violent politics - that is, revolt, terrorism, insurgency, and riot - as "the ultimate method of resolving conflict" which can be avoided only by functioning political institutions (p. 4). I see this as one of the explanations as to why the two strands of literature developed parallel to one another for a long time.

However, empirical, large-N studies on the effectiveness of nonviolence relative to violence were almost non-existent until the influential work of Chenoweth and Stephan (2011). In fact, in their data on mass resistance campaigns, the use of nonviolence is almost twice as likely to yield full or partial success (Chenoweth and Stephan 2011). The monograph is a thorough empirical analysis and present both qualitative and quantitate evidence of why nonviolent campaigns enjoy higher success rates compared to their violent counterparts. For any campaign, attracting large segments of the population is critical in order to inflict higher costs of maintaining the status quo for the government. According to Chenoweth and Stephan (2011), this is the comparative advantage of nonviolent campaigns. To the average participant, protesting may be costly, but violent insurgency is at another level (Chenoweth and Stephan 2011, 37).

While the supremacy of nonviolence is becoming increasingly well-understood, this does not offer an explanation to why some nonviolent protests at some point change to involve the use of violence (Shellman, Levey, and Young 2013). The finding also poses an additional puzzle that this thesis addresses: if nonviolence is both less costly and more efficient than violence, why is it relatively common that nonviolent protests escalate violently?

2.4 What structural theories of violence do not address

Parallel to the study of nonviolent resistance runs a vein of scholarship on violent resistance, typically studied separately from nonviolent tactical choices. For example, there is a large literature on the determinants of specific types of nonroutine violent political action, such as civil war - an extreme form of violent, coercive and contentious political action.

Traditionally, the outbreak of political violence has been explained using largely structural theories, such as the relative deprivation family of theories or theories based on political opportunity structures. Scholars of intrastate political violence, in particular civil war, have focused much attention on country-level, structural factors to explain variations in risk of this extreme form of political dissent (Shellman, Levey, and Young 2013). To exemplify, in the quantitative literature, a range of variables that are associated with an increased risk of violent internal conflict and insurgency have been identified. Examples include poverty and natural-resource dependency (Collier and Hoeffler 2004); mountaneous terrain and state weakness (Fearon and Laitin 2003); regime type and regime change (Gleditsch and Ruggeri 2010; Hegre et al. 2001); and identity-based horizontal inequality and sociopolitical exclusion (Cederman, Gleditsch, and Buhaug 2013; Østby 2013).

Within this research tradition, the country-year has been the dominating level of measurement. The problem with this approach when analyzing shifts that occur *during* a protest is that such structural factors are relatively static and change little over time (Shellman 2006). These models do not take into account the strategic *interaction* between the dissatisfied citizens and the government. First of all, it is problematic, particularly considering the relatively short duration of protest activities, because actors do not respond to the others' behavior at yearly intervals (Davenport 2007).

Second, such macro-level factors are found not to influence whether an individual chooses to participate in nonviolent demonstrations, further motivating the choice to look at situational, rather than structural variables in the forthcoming analysis (Stockemer 2014). Moreover, as shown by Chenoweth, Perkoski, and Kang (2017), traditional structural models fare relatively poorly in predicting the onset of nonviolent protest. This finding indicates that more complex and situational dynamics are at play.

My thesis targets these identified gaps in the literature. Rather than relying on structural theories and explanatory variables, I develop a Rational Actor (RA) model, drawing on the work by Gustafson (2019), Lichbach (1987), and Pinckney (2016). The model accounts for the dynamic interaction between protesters and the government. To this end, I use data that is both spatially and temporally disaggregated.

2.5 Repression and tactical changes

I now turn to one of the explanatory variables that I focus on in this thesis, namely government repression. While the government may not always be the object of the dissent, it is a stakeholder to nonroutine political action because it has the means to supply security forces, to set the rules, or to mediate in the conflict of interest (Tilly 2003).

Facing a dissent group, the government has three distinct choices. First, the government can choose to simply ignore the protest. Second, the state can choose to accommodate the protesters' demands. This relates directly to the magnitude and intensity of the protest (Chenoweth and Stephan 2011). If the costs of maintaining the status quo are greater than the cost of concessions, the government will give in. In October 2015, the demands of the #FMF were quickly met with concessions from both universities and the South African government. When the protests spread throughout the country, the Minister of Higher Education stated that no university fees were to increase with more than six percent, despite the fact that, formally, this is outside the jurisdiction of government (Essop and Nicolaides 2015). A few days later, after violent clashes between protesters and the police outside government offices in Pretoria, President Jacob Zuma declared that there would be no tuition fee increases in 2016 (Pearson, Karimi, and McKenzie 2015). Clearly, the mounting protests had succeeded in increasing the costs of conducting business as usual across the government's threshold, at least in the short term.

Third, the government can choose to repress the dissent. Repression is conventionally defined as coercive measures taken by a state through its security forces against an opposition group, with the aim of increasing the costs of continuing dissent and/or deter specific activities (Davenport 2007; Lichbach 1987; Rasler 1996). "The law of coercive responsiveness" (Davenport 2007, 7) states that governments generally take repressive measures to counter the threat to the status quo that protests represent. Repression can take many forms, varying from applying non-lethal intervention, such as crowd dispersal, to the potentially lethal deployment of security forces commanded to violently supress the protest.

Governments vary in their inclination to use different types of repression (Cunningham and Beaulieu 2010). In autocracies, where the government has total control over the security apparatus, costs of repression are relatively low compared to those of policy compromise (Pierskalla 2010). Democracies and semi-democracies, on the other hand, are thought to prefer compromise to repression. While repression is clearly costly for protesters, there are also considerable potential costs for the government and their security forces. These include, among other factors, audience costs and fear of losing office (Pierskalla 2010, 122; Davenport 2007). Generally, however, states show greater willingness to repress violent rather than nonviolent challengers (Davenport 2007). Not only is repression of violent protest considered more legitimate and justifiable to third parties than repressing nonviolent protest; violent protest may pose a greater threat to the state (Lichbach 1987; Cunningham and Beaulieu 2010).

When reversing the causal arrow and looking at the effects of repression on dissent, there is less consensus. The empirical puzzle that many studies of the repression-dissent nexus seek to explain is the so-called "Punishment Puzzle" (Davenport 2007, 8). That is, that government repression both deters and escalates dissent (Lichbach 1987; Opp and Roehl 1990; Rasler 1996). While there is a vast literature on the subject that is beyond the scope of this thesis to review, little consensus has been found on the mechanisms that lead to the inconsistent findings.

Turning to the influence of repression on the violent escalation of protest, there are also inconsistent findings in the literature. One of the first theoretical contributions to the literature on the impact of repression was made by Lichbach (1987). The author develops an RA model of the interaction between government and protesters, argues that repression deters the particular tactic, which will lead the dissident groups to substitute that method for the other available method. In his model, this means that repression of nonviolent protest leads to an increase in violent protest, while governments that repress violent protest will see an increase in nonviolent protest.

Using event data of dissident-state interaction in Peru and Sri Lanka, Moore (1998) provides empirical support for Lichbach's (1987) so-called substitution hypothesis. Moore's (1998) analysis provide a hard test for the hypothesis that groups substitute nonviolence for violence, and vice versa. This is because the two groups that are studied are described as "especially violent guerrilla movements" and are therefore less likely to switch from violent tactics to nonviolence (Moore 1998, 862). However, the analysis is also limited precisely because it focuses only on two similar cases, although they operated in different political environments. The large, panel dataset that include a myriad of different protests that I use in this thesis will provide additional empirical support for the hypothesis. Furthermore, I provide insight into whether the hypothesis only holds for organized movements, such as Moore's (1998) guerrilla movements, or whether it is generalizable also to relatively disorganized protests. In Chapter 4, I more closely inspect and explain the logic underlying the substitution hyptohesis.

In their paper on two dissident organizations in the Philippines and Sri Lanka, MILF and LTTE, Shellman, Levey, and Young (2013) show that the level of government repression in the previous month is an important predictor of the onset of a violent phase. The authors also find support for the hypothesis that the effect of repression follows an inverted-U shape, so that the risk of violence is highest at moderate levels of repression (Shellman, Levey, and Young 2013). However, similar to the analysis conducted by Moore (1998), this analysis is also limited as it only focuses on two relatively organized and, at times, extremely violent dissident groups. This makes generalizations difficult in two respects: to other types of protest and to other political environments. In sum, scholars have tended to study the effects of repression with a small-N sample, using individual cases or conducting regionally focused studies, which introduces a potential selection bias (S. R. Bell and Murdie 2016).

Gustafson (2019), on the other hand, finds no signficant effect of government repression on violent escalation of nonviolent demonstrations. The author uses a broader sample and lower-level contention than both Moore (1998) and Shellman, Levey, and Young (2013), which brings the external validity of the substitution hypothesis into question. In the analysis, repression is operationalized as the proportion of events that were violently repressed by each country's government in the past year. Similar ways of evaluating the effect of repression has also been used by Pinckney (2016) and Cunningham and Beaulieu (2010). Cunningham and Beaulieu (2010) differentiate between the type of tactic (nonviolent or violent) that was repressed using a monthly rolling average. Their results show that repression of violent dissent decreases the use of violent tactics relative to nonviolence, in line with the substitution hypothesis. Similarly, governments that repress nonviolent strategies are likely to see increasing relative use of violence.

In his monograph on the sustenation of nonviolent discipline in civil resistance campaigns, Pinckney (2016), too, finds that repression of nonviolence is a substantively important and statistically significant predictor of violent events. The author operationalizes repression similarly to Gustafson (2019) and Cunningham and Beaulieu (2010) as the occurrence of repression in the recent past, defined as the average number of the campaign's past 25 actions that were repressed. In contrast to Cunningham and Beaulieu (2010), Pinckney (2016) does not find that government repression of violent events influences tactical choices in any direction.

To summarize, most studies that have assessed the influence of repression on tactical choices have operationalized this as the level of repression in "the recent past". One sound reason to this operationalization is to avoid problems of endogeneity and reverse causality (Gustafson 2019). Without sufficiently disaggregated data, one cannot infer that repression led to violent escalation and not vice versa. It is also commonly assumed that protesters estimate the likelihood of repression and its probable intensity based on previous experiences, and choose tactics on the basis of this (Cunningham and Beaulieu 2010). Comparatively less research has been made on the more immediate impact of repression on tactical choices. In his quantitative analysis, Pinckney (2016) states that he did not find any statistically significant effect of repression of one event on the likelihood that the event that followed immediately after was nonviolent (p. 39). Unfortunately, the author does not report the regression results that support this theoretically important finding. In my analysis, I look at whether repression in week$_{t-1}$ leads to tactical shifts in week$_t$. In other words, instead of "the recent past", I look at "the immediate past" which has been understudied.

2.6 Organization and tactical changes

repression of nonviolent protest can lead to so-called *political jiu-jitsu* or backfire (Sutton, Butcher, and Svensson 2014). This occurs when government's use of power against unarmed protesters reduces its legitimacy and power, which may lead to heightened protest frequence or intensity, increased sympathy for the protesters and mobilization, or even a spiral of escalating violence (Opp and Roehl 1990; Schock 2013; Sutton, Butcher, and Svensson 2014). In order to generate backfire, studies have highlighted the importance of effectively communicating the repression to relevant audiences through organized communication structures (Sutton, Butcher, and Svensson 2014), as well as protesters sticking to nonviolent tactics even when facing brutality (Schock 2013). Nonviolent discipline may therefore be viewed as a requisite for successful nonviolent protests.

In 2011, what started out as peaceful demonstrations developed into a series of revolutions and protests, both violent and nonviolent, that quickly spread across the Middle East and North Africa (MENA) region. In Egypt and Tunisia, the protests were loosely organized, working in the favor of the protesters as the two regimes did not allow for any organized opposition (Kaphahn and Brennan 2017).

This did not, however, preclude the violent escalation of protests.

As mentioned, much of the literature on protest uses the campaign as the unit of analysis. Theoretically, where a leadership can be identified, it is frequently assumed that the group's tactics are decided upon by the leader. In his study on the effect of government repression on dissident tactics, Lichbach (1987) defines dissident groups as teams, characterized by a shared, overarching goal. In other words, in the context of protest, one individual participant's goal is identical to the others'. Individual protesters are considered rational actors, and as they agree upon the group-level goal (e.g. regime change), the group's goal "can be viewed as a single consistent preference ordering" (Downs 1957, cited in Lichbach 1987, p. 278). In both conceptualizations, the protest actor is assumed to be a unitary actor.

Although assuming actor unitarity simplifies the formal model, it also overly simplifies reality (Pearlman 2010; Wilkinson 2009). As noted by Chenoweth and Stephan (2011), the use of violence in protest movements is often initiated by fringe factions that defy or act independently of the leadership. However, simply stating that these violent individuals are irrational would be an overly simple resort (Gustafson 2019; Lichbach 1987; Wilkinson 2009). Indeed, movement fragmentation is argued to increase the risk of protest violence, because protests coordinated by organized groups are more able to enforce and articulate nonviolent discipline and a cohesive strategy (Pearlman 2012). Thus, actor unitarity may be more likely to be observed in organized protests, while individual preferences shape the tactics of disorganized groups. The theoretical model in this thesis is therefore based on individual preferences, which are constrained by the organization's preferences, if there is a level of organizational capacity behind the demonstration (Gustafson 2019). Therefore, organization is treated as a variable rather than an assumption behind the unit of analysis.

The impact of organization on maintaining nonviolent discipline is however not straightforward. In his quantitative analysis, Pinckney (2016) does not find the expected, positive relationship between the strength and cohesion of campaign leadership and ability to maintain nonviolent discipline when controlling for repression.

Yet, because the author uses the campaign as the central behavioral unit, this finding does not really tell the story of unorganized protests.

Throughout the last decade, a new wave of protests have spread across Africa. Many of these have been characterized as dissimilar to the 2011 Arab Spring, insofar as they generally did not have revolutionary intent (Harris and Hern 2018). Instead, Harris and Hern (2018) argue that in Africa, taking to the streets is a channel to express political preferences and demands for e.g. better public services or material goods. In the forthcoming analysis, I use protest event data from Africa and some countries in Asia from 2010 to 2018. Thus, according to the conceptualization of Harris and Hern (2018), a great deal of these contemporary protests lack the ideological cohesion that characterize social movements. This further motivates my choice to move away from the campaign as the unit of analysis.

Therefore, I take on a different perspective. As campaigns are defined as behavior that is sustained over time by a (single) named group, they are conceptually separated from more spontaneous events or events of shorter duration (Chenoweth and Stephan 2010). In my analysis, I do not distinguish between long-term and short-term demonstrations[2], thereby recognizing that violent escalation may in fact be more likely in spontaneous and less-organized events. In other words, I look at *both* organized and less-organized protests, something that has largely been overlooked by scholars on nonviolent and violent resistance alike. Gustafson (2019) is a prominent exception; finding that spontaneous demonstrations have a higher risk of violent escalation. I therefore provide additional empirical evaluation of the relationship between organization and tactical changes in protest.

2.7 Knowlegde gaps: Repression and organization

In this section, I provide a brief summary of the knowledge gaps identified in the above sections along with how I address these in my thesis.

First, while there is an abundance of literature on the onset of violent and nonvio-

[2]However, protest cycles are required to have a temporal span of two weeks or more, in order to measure change

lent dissent, comparatively less systematic inquiry into tactical changes during the course of *demonstrations* has been made. Gustafson (2019) is a notable exception. Yet his study is limited in two key ways. First, the author models only unidirectional escalation from nonviolent demonstration to violent riots using a binary dependent variable. Second, his analysis does not take into account the dynamic interaction leading up to tactical changes. I build on the studies of Gustafson (2019) (who study protest events), Pinckney (2016) (who study campaigns), and Shellman, Levey, and Young (2013) (who study violent organizations), and offer a disaggregated theoretical model and research design. The focus is more on situational variation rather than structural variables, which has long been the norm in quantitative studies of political violence. I specifically look at weekly changes in the level of violent protest, allowing me to assess the course of events leading to eventual escalation.

Second, and relatedly, most analyses of nonviolent protest and tactical changes and the effects of repression are situated on the campaign- or organization-level (Chenoweth and Stephan 2011; Pinckney 2016). The definition of the campaign includes that there is a certain level of organization to the protest. Thus, the spontaneous and disorganized sparks that may eventually cause a prairie fire are, by definition, excluded from much of the scholarly work on protest. In this thesis, I offer an alternative approach. Rather than analyzing protest as campaigns, and thereby implicitly assuming actor unity, I look at the total protest activity at a specific point in time and space. I explore whether there exists a link between the level of organization "on the ground" and the level of violence employed by the protesters. This is an aspect largely left unexplored in the literature due to the focus on campaigns. My data provides a hard test for this hypothesized relationship, as I do not confine the analysis to campaigns where there is by defintion an existing organizational structure.

Finally, to my knowledge, this thesis is the first to explicitly address the potential interaction between level of organization and repression that is not on the campaign level of analysis. A common assumption in the literature on nonviolent resistance is that government repression is less effective against highly organized movements

(Chenoweth, Perkoski, and Kang 2017). However, there is a dearth of empirical literature that explicitly addresses this by treating organization presence as a variable. In my theoretical model, I argue that the influence of repression on tactical changes cannot be analyzed in isolation from the level of organization. While Lichbach's (1987) substitution hypothesis has received much empirical support (Cunningham and Beaulieu 2010; Moore 1998; Pinckney 2016; Shellman, Levey, and Young 2013), I examine the external validity of the the hypothesis by using a large sample and treating repression as a variable that is conditional upon organization.

3 The dynamics of protest events: The Rational Actor model

In this chapter, I present the theoretical framework that I have developed for this thesis along with the hypotheses. To theorize the relationship between repression, organization and tactical changes, I turn to the Rational Actor (RA) model. I begin by explaining why I have chosen the RA model. To lay out the theory, I specify four critical concepts of the RA model in the context of protest politics: The actors and their available tactics, the goals (benefits), and the costs associated with the tactics (Lichbach 1987). The mathematical models are adopted from Pinckney (2016) and Lichbach (1987). In his monograph, Pinckney (2016) models the variation in the ability of anti-government campaigns to uphold nonviolent discipline in non-democracies. This model has also been applied by Gustafson (2019), who more explicitly focused on the effect of individual-level grievances on the violent escalation of nonviolent protest. Lichbach (1987) presents a mathematical model that accounts for the various effects of repression on dissident tactics.

The hypotheses I derive are founded on two key variables that I argue influence the costs of continuing nonviolent protest, thereby influencing the risk of violent escalation. The first set of hypotheses are drawn from Lichbach (1987), and concerns the influence of government repression. The second hypothesis relates to the level of organization on the ground, which is argued to positively influence nonviolent discipline. Finally, I argue for an interaction between repression and organization.

3.1 Why rationality?

One of the main advantages with applying the RA model is that it provides a sufficiently general framework for explaining behavioral choices grounded in theoretical expectations as well as allowing for the prediction of behavioral shifts in the short term (Riker 1995). Assuming that protesters are rational and choose tactics based on an assessment of costs and benefits resulting from their available choices is the most widely used model in the literature on tactical choices, providing a base assumptions to the theory that is both parsimoneous and intuitive.

The core assumption made in the RA model is that an actor have ordered preferences or goals (Riker 1995). In order to reach her most highly valued goal, the actor will choose the "best" means (Riker 1995). In other words, the actor decides on specific actions (i.e. tactics) on the basis of a comparison of the expected costs and benefits across the available choices (Cunningham and Beaulieu 2010, 174). Actors seek to minimize their costs while at the same time maximize their benefits. In this sense, the concept of *rationality* means that actors purposefully choose means that they believe will produce their most preferred outcome with the least costs (Cunningham 2013; Riker 1995). While it may be argued that this is an oversimplification of reality (for a more nuanced approach, see Pearlman (2010)), the model is widely used in the literature to study the dynamic interactions between dissidents and their opponent (see e.g. Gustafson 2019; Lichbach 1987; Pinckney 2016). Another widely used approach involves theorizing the interaction between conflictual parties as a bargaining process, however this is unsuited for the analysis of disorganized protests (for a more thorough discussion of this, see (Gustafson 2019, 5–6)).

In a dynamic RA model, the relative costs incurred, and the benefits reaped from different choices may be assumed to change during the course of events dependent on the choices made by other actors, as "actions by one actor produce outcomes in conjunction with actions by others" (Riker 1995, 25). The assessment of utility is therefore considered as a continuous process of estimation and re-estimation. The strategic nature of the interactive behavior between conflictual parties is the core of the RA model described in this chapter.

Modelling the strategic interaction of conflictual parties is particularly well-suited in the study of protest as these are relatively short-lived phenomena that often involve some form of interaction between the protesters and the target of the protest (Shellman, Levey, and Young 2013). In order to explain the violent escalation that sometimes occur within this short time frame, a temporally disaggregated and dynamic theoretical framework is essential. This is best provided by the RA model

(Lichbach 1987). While traditional, structural approaches have been successful in unveiling general patterns of political violence, assessing the influence of structural and largely static factors is futile when change occurs within the span of days or weeks.

At onset, it is safe to assume that engaging in nonviolent protest is less costly than violent protest for the individual participant, simply because it is less personally dangerous (Chenoweth and Stephan 2011). Generally, the prospects of success, i.e. the end-state benefit, are greater for protesters that remain nonviolent (Chenoweth and Stephan 2011). Protesters choose to rely on the tactic that is most effective at producing government concessions (Moore 1998, 853). According to Eckstein (1980), the occurrence of collective political violence is the result of a temporal process during which "the *costs* of violent collective action are expected to be especially low" or "nonviolent actions in pursuit of highly valued goals have been shown to be unproductive" (Eckstein 1980, 155, emphasis added). The use of violence is therefore the result of strategic calculation made by a collective of individuals, as is the use of nonviolent action (Eckstein 1980, 147). In other words, while nonviolence may often be the preferred tactic at onset, if the prospects of success by using violence increases, the protesters may prefer to switch tactics (Gustafson 2019).

The model applies to situations where a group of protesters have already taken to the streets. At this point, the total protest activity is not 100 percent violent, as I am interested in what causes violent escalation rather than the onset of violent protest. The presumably disaffected group of citizens have thus already overcome the collective action problem of mobilization and organization (Chenoweth and Stephan 2011). This implies that the personal costs of protesting in the first place are outweighed by both the expected benefit and probability of success. Furthermore, they have chosen protest as their preferred mode of expressing dissent. I therefore do not propose a model of the choice of violence over nonviolence at onset; nor do I make predictions about whether any given tactic will be successful.

3.2 Actors

The key actor that I consider is the individual who has taken to the street. To keep the model parsimoneous, I assume that the two tactics that the protester can choose from are nonviolent protest and violent protest. These are distinguished from one another simply by the absence or presence of intended physical harm toward people or property (Bond et al. 1997). Protesters are further assumed to be both willing and able to substitute nonviolence with violence, and vice versa (Moore 1998, 853). This follows directly from the above assumption that protesters are interest-maximizing actors seeking to reach some predefined policy goal.

The observed behavior that I study is thus the sum of individual decisions to refrain from or engage in violent behavior (Pinckney 2016). Individual protesters' preferences can be modified by whether the protest is organized or not, as it is assumed that an organized nonviolent protest has a protest-level preference to remain nonviolent (Gustafson 2019, 6). I thereby treat the organization level as a variable, rather than as an assumption underlying the unit of analysis. Specifically, I look at the proportion of violent protest activity relative to the total protest activity. This relative frequency depends on the choices of a relatively large number of individual protesters to use violent means. The assumption that the observed behavior is a result of strategic and rational calculation allows for modeling behavioral shifts.

Protest actions do not occur in a vacuum. Because protest is a channel for political action, I also consider interactions in the broader political environment in which protest occurs. This involves defining a second actor, namely the government. Governmental bodies, such as police and other public security forces are assumed to act as agents of the government. The two actors, the government and the protesters, do not act independently of one another (Riker 1995). For the sake of simplification, I assume that they act independently of third-parties. The government is assumed to want to minimize dissent (Davenport 2007).

3.3 Modelling tactical choices

I now turn to present the formal model that I apply to explain shifts from nonviolence to violence.

Pinckney (2016) argues that if the following conditions hold, the protester will prefer to use nonviolent direct action over violent resistance:

$$B_{nv} - P_{nv}(R_{nv}) + P_{nv}(S_{nv}) > B_v - P_v(R_v) + P_v(S_v)$$

The terms in the model are defined as follows (Pinckney 2016, 23):

- B_x is the benefit associated with a given tactic x, i.e. nonviolence or violence.
- P_x is the probability of any punishment or reward from using a given tactic x.
- R_x is the expected intensity of government repression of a given tactic x.
- S_x is the expected intensity of any punishment or reward by the movement for using a given tactic x.

Intuitively, the protester will choose to act in a nonviolent manner if the (perceived) benefits of doing so are larger than those attained by using violence. Here, there are two sets of benefits to be considered. First, individuals have a set of prior and personal preferences for nonviolence or violence (Pinckney 2016). These are shaped by both personality traits as well as experience and training. This is one source of unobserved differences between individual protesters.

Second, the benefit associated with an action can be conceived of as the achievement the protest demands, such as obtaining policy concessions by the government (Lichbach 1987). If the end-state benefit is most likely to be achieved through the use of nonviolence, as statistically shown by Chenoweth and Stephan (2011) to be the case, the rational choice for the protesters would be to use nonviolent protest. This implies that external factors, i.e. punishments, rewards or repression, are likely to influence the choice of using violence once the protest has begun.

Nonviolent protests may escalate violently in two situations, as described by Eckstein (1980) in the beginning of this chapter. First, if the costs of violence are lowered relative to the costs of nonviolence. Second, if the perceived benefit derived from using violent tactics increase relative to the benefit of remaining nonviolent. An added factor from the model above is whether there is a protest-level capacity to enforce nonviolent discipline. Both of these situations can materialize if one individual chooses to protest violently, and thereby lowers the costs of violence for the other participants and increases the efficacy of violence (Gustafson 2019, 6).

In this section, I present two factors that are argued to influence this cost-benefit calculation: government repression and level of organization. I focus on government repression because, according to Pinckney's (2016) model, if nonviolent protesters assume that they will be repressed, this influences whether they will remain nonviolent or not. The level of organization on the ground also influences this calculation. If the protest, whether nonviolent or violent, is unorganized or spontaneous, the term $P_x(S_x)$ equals zero. It is often assumed that highly organized protests lead to a higher likelihood of protests remaining nonviolent; however this is largely left unexplored in the quantitative literature.

3.3.1 Government repression

The protesters' costs related to tactical choices are in large part determined by the state's response (Cunningham and Beaulieu 2010; Lichbach 1987). The cost of being repressed is represented by two terms in Pinckney's (2016) model. The protester expects nonviolence to be repressed with a probability P_{nv} and with an intensity R_{nv}. Repression of violent protest is expected with a probability P_v at an expected intensity of R_v. If we assume that $B_{nv} > B_v$, and that there are no movement-level punishments or rewards, the equation above suggests that government repression of nonviolence must be greater, both in terms of probability and intensity, for nonviolent tactics relative to violent tactics if the benefits of nonviolence are to be outweighed by violence. In simpler terms, for protesters to choose violence over nonviolence, they must expect that nonviolence will be repressed and that this repression will be more severe than if they use violence.

Lichbach (1987) presents a parsimonious RA model that provides an answer to the aforementioned Punishment Puzzle. In the model, the protesters face a choice between engaging in nonviolent or violent activities. The central question in his article is how the total amount of protest activity changes in response to changes in government repression (Lichbach 1987, 275). Total protest activity is modeled as the sum of nonviolent and violent activities. I focus on how his models explain changes in the relative amount of violence to nonviolence.

For the protesters, there are specific costs associated with both tactics, given by the cost function:

$$C = nT_n + vT_v + F \text{ (Lichbach 1987, 276).}$$

C denotes the total costs, n is the cost per unit of nonviolent activity, T_n is the total amount of nonviolent activity, v is the cost per unit of violent activity, T_v represents the total amount of violent activity, and F denotes fixed costs associated with overcoming the collective action problem. Given that protesters are rational and cost-minimizing actors, they thus choose the relatively cheapest tactic. In Pinckney (2016)'s model, the costs per unit of each tactic in terms of repression is denoted by $P_{nv}(R_{nv})$ and $P_v(R_v)$.

Lichbach (1987) proceeds to demonstrate mathematically that the protesters' relative costs related to a given tactic changes in the face of repression. The direct costs of a given tactic, n and v, are products of the severity of government repression of T_n and T_v, respectively. In other words, repression is costly. The protester's "budget constraint" is therefore "the amount of repression [she] is willing to incur to obtain the government ouput [she] desires" (Lichbach 1987, 277). These costs can be thought of as the direct results of the repressive measures taken by the government, such as injuries from clashes with police forces. Because protest groups choose tactics based on the assessment of costs and benefits, these changes in relative costs are important to take into consideration when modeling the shift from nonviolence to violence.

The logic that underlies the so-called substitution hypothesis is as follows: If the government represses, rather than accommodates, nonviolent protest, it increases the cost of nonviolent protest relative to violent protest; a conclusion that follows from "the assumption that repression is costly" (Moore 1998, 854). Vice versa, if the government represses violent protest, the costs of violent protest increases relative to nonviolent protest. If the costs of nonviolence are sufficiently great relative to violence, the initial assumed preference of nonviolence over violence, $B_{nv} > B_v$, may be offset. In other words, because protesters are interest-maximizing actors, and if the government responds to nonviolent protest with repression, the protesters will shift from nonviolent to violent tactics, and vice versa (Moore 1998, 853). Thus, an increase in repression of nonviolent protest leads to a reduction of this form of protest, but an increase of violent protest (Lichbach 1987, 285).

During the #FMF protests, students alleged that they turned to violent means after police forces were sent to campuses in an effort to disperse the protests as this silenced the ongoing nonviolent protest (Nicolson 2016). The repressive response signals important information about the regime: by dispatching security forces to contain protests, it proves both able and willing to repress unarmed dissidents, which can induce additional mobilization (Sutton, Butcher, and Svensson 2014). In addition, it can lead to further perceptions of unjust and anger. In this situation, nonviolence has proven ineffective. If the costs of nonviolence are sufficiently great relative to violence, the initial assumed preference of nonviolence over violence, $B_{nv} > B_v$, may be offset. Similarly, if a protester learns that violence is met with violence, she may attempt the different method to increase the likelihood of success in the end. This does not only pertain to what the individuals experience themselves, but also what is observed in their surroundings, because they expect consistent responses from the government[3].

Now, recall the conditions proposed by Pinckney (2016):

$$B_{nv} - P_{nv}(R_{nv}) + P_{nv}(S_{nv}) > B_v - P_v(R_v) + P_v(S_v)$$

[3]The importance of consistent government responses of accommodation or repression is also modeled by Lichbach (1987)

To summarize, nonviolent protesters will remain nonviolent if the costs of doing so are lower than for violence. In the opposite direction, a rational protester will choose violence over nonviolence if the costs of doing so are lower than for nonviolence at a given point in time. The costs of protesting are in large part determined by the state's response. Following the substitution hypothesis, repression of nonviolence increases the costs of nonviolent protest relative to violent protest (Lichbach 1987). This holds also in the opposite direction: repression of violence increases the costs of violent protest relative to nonviolent protest.

The probabilities of each form occurring, P_{nv} and P_v, is captured by introducing a one-week lag of the repression variable. If nonviolent protest in week$_{t-1}$ is repressed, the protesters are assumed to have observed this and expect a higher probability of repression of nonviolence in week$_t$. Again, the same is assumed to hold for violent protest, so that repression of violent protest is expected with a higher probability in the subsequent week.

Moreover, given that nonviolence is the most effective tactic for the protesters to achieve their policy goal, government repression of nonviolent protest may in fact heighten the overall level of conflict, $T_c = T_n + T_v$ (Lichbach 1987, 285). Assume that nonviolence is twice as effective as violence in terms of producing the desired output and that the government chooses to repress the nonviolent protest. According to the model, repression will then not only reduce the level of nonviolent protest, it will also increase the level of violent protest. For the protesters to achieve success, they will have to substitute twice the amount of the less-effective tactic to compensate for the decrease in nonviolent protest activity (Lichbach 1987, 286). Thus, repressing nonviolence may lead to a spiral of political violence. This theoretical implication highlights the policy-relevance of studying various government responses to protest.

Apart from the influence that repression has on the cost-benefit calculation, repression can also lead to violent escalation through other mechanisms. Repression of nonviolent protest can increase anger or fear among the protesters, which may in turn lead them to take to violent means in order to protect themselves. Nonviolent discipline is less likely to be observed where the state uses repression (Pinckney

2016). With regards to repression of violent protest, repression may be more intense as it is considered more legitimate to repress violent protesters (Davenport 2007). This may effectively quell protest, not only by raising the costs of continuing violent dissent, but also by the show of superior force. These alternative mechanisms have the same implications with regards to violent escalation as the one accounted for by the RA model. Some may contend that repression of violent protest can lead the protesters to continue or escalate their use of violence as a means to protect themselves. This is not captured by the RA model due to the underlying assumption that repression is costly, and that protesters choose the relatively cheapest tactic.

3.3.2 Level of organization

If a nonviolent movement is able to efficiently punish individual's use of violence, or effectively enforce nonviolent discipline, the risk of violent escalation decreases (Pinckney 2016). Again, I refer to the mathematical model:

$$B_{nv} - P_{nv}(R_{nv}) + P_{nv}(S_{nv}) > B_v - P_v(R_v) + P_v(S_v)$$

The terms S_{nv} and S_v describes the possibility of the movement to respond to the tactical choices of their members using rewards or punishment, with a probability of P_{nv} and P_v.

If $P_v(S_v)$ increases, i.e. if the protester sees it as likely that violence will be punished in some form, the relative costs of violence increases. Similarly, obtaining rewards for remaining nonviolent, which can be conceptualized simply as positive affirmation from fellow protesters, increases both the relative benefit of nonviolence as well as the costs of turning to violence (Pinckney 2016).

Any actual punishments or rewards to individual protesters by a movement, such as peer affirmation, are not directly observable. However, whether the protests are organized or not, is. The costs of continuing nonviolent protest is lower where there are "organizational capacities" behind the protest (Pierskalla 2010, 119). Protests that are organized decrease the costs of remaining nonviolent through at least two mechanisms. First, where there is an organization behind the protest, the leadership can credibly signal to the protesters that the nonviolent action will continue, constraining individual decisions to turn violent in a desperate manner (Gustafson 2019). This guarantee is less present in disorganized or spontaneous demonstrations. In Pinckney (2016)'s terms, these signals can come in the form of punishments or rewards. Second, where there is a pre-existing organizational capacity to the protest, the leadership is more likely to be able to effectively communicate tactical choices and enforce nonviolent discipline in the face of repression (Sutton, Butcher, and Svensson 2014).

Not only can organized protests potentially punish the protesters, and thereby increase their relative costs of resorting to violence; organization makes nonviolence more beneficial in itself. In coordinated protests, the underlying organizational structures are more likely to facilitate important "in-group norms of solidarity and obligation" (Sutton, Butcher, and Svensson 2014, 563), and thereby enforce cohesive strategies of contention (Pearlman 2012). Furthermore, organization "subjects individual decision-making to the constraints of leadership, organizations, and an overriding sense of collective purpose" (Pearlman 2012, 30). It is therefore assumed that there is a level of unity in organized protests that discourages individuals resorting to violence.

To illustrate this relationship, I again return to the #FMF protests. Despite a multitude of underlying causes, the #FMF protest actions were a spontaneous reaction to the fee increases. In the early days of the movement, there is little evidence of a central organizational body that coordinated the activity and had the means to respond to the tactical choices of the students with neither reward nor punishment. This exemplifies a situation where a majority of the protesters want to remain nonviolent, while a minority is becoming increasingly angered and seek violent means. If the protest is unorganized and there is no clear leadership to enforce the nonviolent strategy, the situation may quickly become chaotic, as it did on several South African campuses in 2015 and 2016.

I not only expect that the level of organization has an independent effect on the risk of violence. In addition, if government repression is less effective against highly organized protest, I expect there to be an interaction effect between the proportion of organized protests and repression. The substitution hypothesis suggests that shifts from nonviolence to violence occurs when the former is being repressed, and vice versa. Based on the assumption that repression of organized groups is less effective, I expect that the substitution effect of repression is conditional upon the level of organization. Given that repression lowers nonviolent discipline, organizations can play an important role in reinforcing it (Pinckney 2016). More specifically, I hypothesize that for nonviolent protesters, the costs incurred by repression, that influences tactical shifts, are at least to some extent alleviated by the level of organization. If there is a high level of protest organization, the costs of switching tactics from nonviolence to violence may in fact increase, even if repression occurred in the previous week, which influences their perceived likelihood of current repression.

The association between organization and repression of violent protest cannot be viewed in isolation either. Where violent escalation has already occurred and violent protesters are faced with repressive measures, the costs of continuing violence are expected to increase more where organizational bodies that coordinate nonviolent protests exist, through the abovementioned mechanisms. In other words, I expect that the substitution effect is more pronounced in situations where there are higher levels of organization. For disorganized protests, the reverse relationship is expected also for repression of violent protests, leading to increasing violence.

3.4 Hypotheses

To summarize the theoretical model laid out above, I now present the hypotheses that will be tested in the following analysis.

Following Lichbach (1987), the relative costs of nonviolence to violence are influenced by government repression. This leads protesters to choose the relatively cheapest tactic. The first set of hypotheses are thus:

$H_{1.1}$: Government repression of nonviolent protest leads to violent esca-

lation.

> *$H_{1.2}$: Government repression of violent protest leads to violent de-escalation.*

Second, I expect that organized protests constrain individuals from resorting to violence, both by making nonviolence more beneficial in itself, increasing the costs of violence, and clearly signaling commitment to the nonviolent cause:

> *H_2: Disorganized protest activity leads to violent escalation.*

Third, I hypothesize that repression of organized protests is less likely to lead to subsequent violent protest, while repression of disorganized protest is more likely to escalate violently. The final set of hypotheses to be tested in this thesis are therefore as follows:

> *H_3: Higher levels of organization will moderate the effect of repression on violent escalation.*

4 Data and research design

Analyzing the dynamic interactions between protesters and their opponent requires a unit of analysis that is not the standard country-year or conflict-year approach that has dominated the field of quantitative research on peace and conflict (Davenport 2007; Hegre et al. 2017, 115). Event data, in which the unit of analysis is the event rather than the country or protest movement, is therefore a natural point of departure. In this chapter, I begin by describing in detail the data source and the trade-offs made when using events data. I proceed by explaining the PRIO-GRID unit of analysis chosen for the analysis, which is the total protest activity at a given point in time and space. Finally, I turn to the operationalization of dependent, independent and control variables.

4.1 Data source: ACLED

For this study, I have chosen the Armed Conflict Location and Event Data Project (ACLED) for my coding of the dependent and independent variables (Raleigh et al. 2010). ACLED tracks, among other events, occurrences of political violence and demonstrations on a day-to-day basis. The current version of the data set includes events in 93 countries across Africa, Asia, Eastern Europe and the Middle East.

The unit of analysis in ACLED is the unique event happening on a specific day at a specific location. During the #FMF protests, there was protest activity at several different campuses across the country. Between these, there was considerable variation in how the protests developed. As previously mentioned, at some locations, the protests died out, while at other places, they "erupted into mayhem" (Poggi 2015). To account for subnational variation, I use the temporal and spatial information coded for each protest event. Instead of treating protests as campaigns, I look at where and when protests occurred. I therefore assume that the individual protester is more likely to be influenced by the situational circumstances "on the ground", at the point in time and space where she finds herself. I contend that this is beneficial in order to model the violent escalation and de-escalation of protests.

In ACLED, each event is classified according to its type and by the interacting actors (if any). Because my research question concerns why some protests escalate violently while others do not, I have limited the data to include only demonstration events which are coded as either peaceful or violent. Nonviolent protests includes two subtypes: nonviolent protests with intervention (mild repression of nonviolent protest) and excessive force against nonviolent protesters (harsh repression of nonviolent protest). Here, there is an observed interaction between the protesters and a second actor, which I have limited to include only state-controlled bodies, such as police or military forces. In ACLED, the same categorization is not used for violent protests, and I therefore manually coded repression of violent protest, which I return to below. As the data includes five distinct types of protest, it is sufficiently disaggregated with respect to event type.

Table 1 summarizes the total number of unique events in the sample. These are used to operationalize the dependent and independent variables below. Nonviolent and violent protest are the general types and therefore include events of repression. Thus, of the roughly 44 000 nonviolent protests, 2 110 of these were met with mild repression. Of the 7 037 violent protests, 3 266 were repressed.

Table 1: Frequency table, events of interest

Event	Frequency
Nonviolent protest (NVP)	44101
Violent protest (VP)	7037
Mild repression NVP	2110
Harsh repression NVP	693
Repression VP	3266

Within the data, the temporal coverage varies considerably. The data is continuously collected and released, allowing for a near real-time analysis. However, because I am drawing on data from other sources that are updated on a yearly basis to operationalize the control variables, the cutoff is set to 2018. Events in the European countries included are recorded only from 2018; events in the Middle East are coded from 2016; in Asia, the starting year varies from 2010 to 2018; while events in Africa are most comprehensively coded from 1997. Due to concerns of consistency, I have chosen to include only events in Africa[4] and the countries in Asia that are coded from 2010 onwards in my analysis, thereby excluding the other regions (a full list of all countries included in the sample is found in Appendix). Choosing this cutoff allows me to investigate protests in a wide variety of polities during the past decade, an approach that has been rare in existing studies of protest escalation.

To illustrate the main reason to why I chose 2010, Figure 2 shows the yearly distribution of the relevant types of events in Africa that are reported by ACLED. Most striking is the large increase in the overall frequency of events occurred from 2010 to 2011. After 2011, the number of protest events has remained at very high levels compared to the preceding decade, and continues to increase. We also see an increase in the number of violent events.

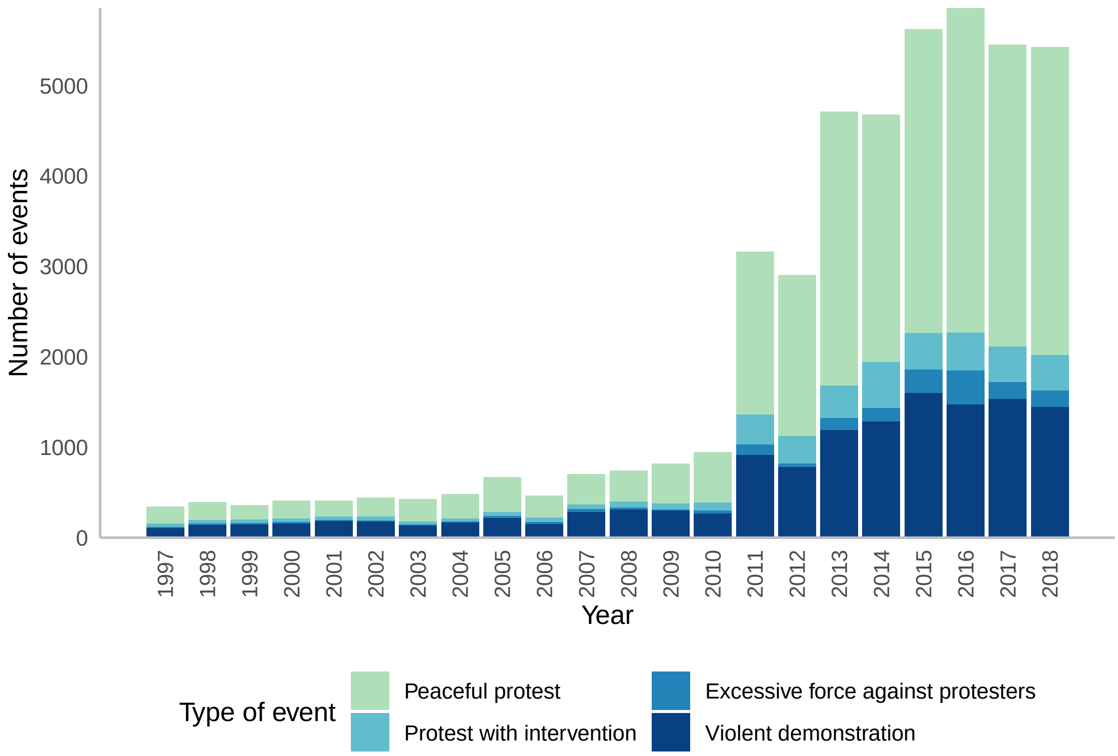

Figure 2: Total protest events in Africa, 1997-2018 (ACLED)

Part of the initial increase can undeniably be attributed to the 2011 Arab Spring in

[4]The countries in Africa that are not included in the sample (Equatorial Guinea, and Eritrea) did have recorded events during the time period, but none that conformed to my definition of protest cycles elaborated upon below.

Northern Africa[5]. However, the Arab Spring cannot alone account for continuity of high frequencies post-2011. According to a report published by the African Development Bank, both the frequency and intensity of civil protest and strikes in Africa have increased during the course of the 21st century (*African Economic Outlook 2017* 2017, 131).

However, the spike in number of reported events can also be attributed to the fact that the current version of ACLED was released in 2010. Events coded prior to this can necessarily only have been coded by tracing media reports back, while events that have happened after 2010 are coded in real-time. This applies also to events in the Asian countries that are included in the dataset from 2010 onwards. Thus, it is possible that only events attracting extensive media coverage are included in the pre-2010 data, creating a possible source of sampling bias. Furthermore, because of the spread of the Internet, I consider it to be more likely that also smaller events have been reported in the 2010s.

Figures 3 and 4 show the total number of protest weeks by country in Africa and Asia, respectively. From figure 4, it may seem striking that India is not included in the sample. This is because India was not included in ACLED before 2016.

4.1.1 Reliability of event data

The availability and use of fine-grained event data in peace research has surged as a consequence of the development of machine-learning techniques and automatization efforts that have considerably eased the workload of the scientists that develop these datasets. This development is however not without pitfalls. The automatization of data generation may lead to a lack of nuance and detail that are products of human interpretation. Many of these large datasets draw on media reports.

Chenoweth, Perkoski, and Kang (2017) notes four biases that can be present in media-based data: "an urban bias, a violence bias, an English-language bias, and a state-centric bias" (p. 1956). Weidmann (2015) links the well-known issue of selective

[5] This is also why I chose 2010, the year the current version of ACLED was released, rather than 2011 as my cutoff.

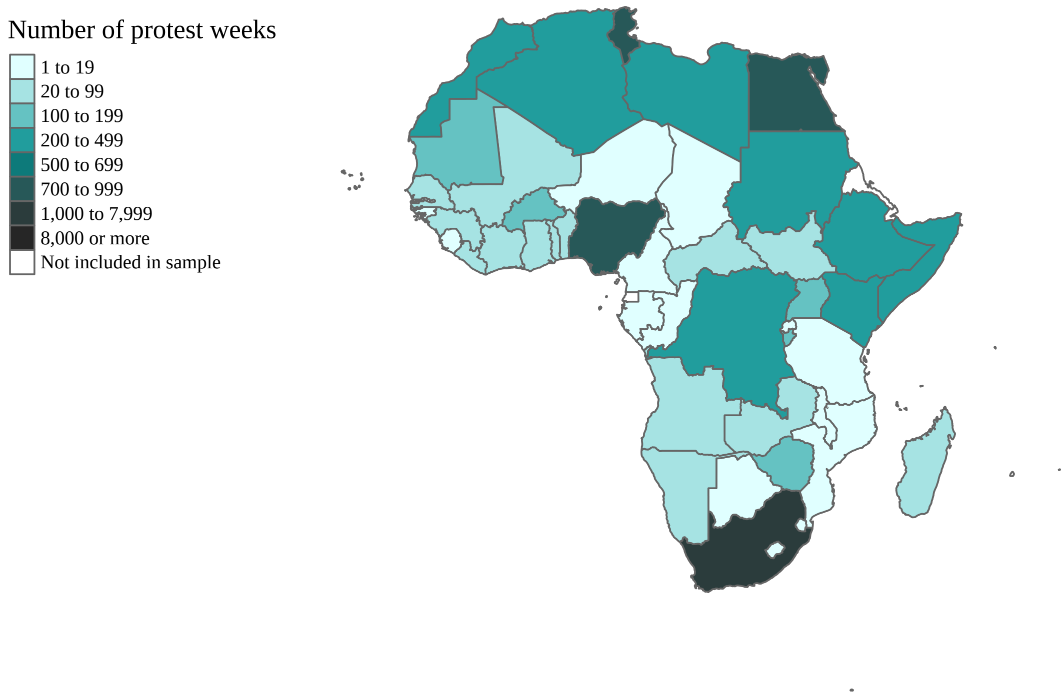

Figure 3: Number of protest weeks in Africa, country level (ACLED)

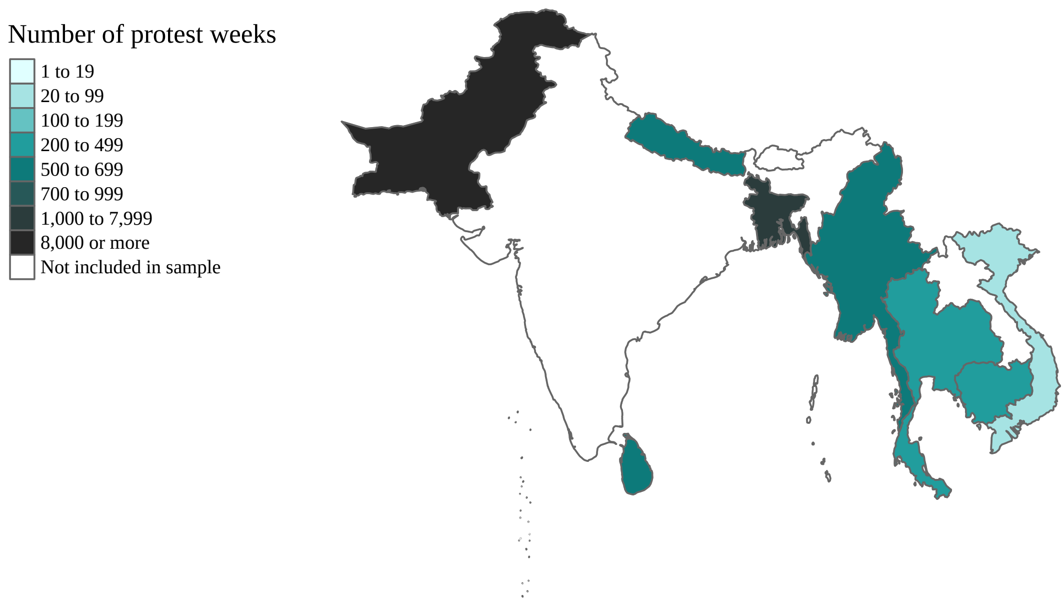

Figure 4: Number of protest weeks in Asia, country level (ACLED)

reporting of sensational events to measurement error. That is, the underreporting of less 'salable' events which leads to a potential problem of inference (Weidmann 2015, 208). This is true too for protest events, particularly those that do *not* involve violence, rioting and police brutality. In other words, protests may be covered by the press once they turn violent, while preceding and parallel nonviolent action remains unreported. Relatedly, nonviolence or small-scale violence in protest settings may not be reported by media outlets compared to large-scale violence, leading to an error of measurement on the dependent variable.

Statistical problems arise if the measurement error correlates with the independent variable(s), which may lead to biased estimates both in terms of direction and magnitude (Weidmann 2015, 208). Ususally, however, studies assume that the measurement error is unsystematic which only leads to increased uncertainty of estimates, which in turn increases the risk of type-II error (Weidmann 2015, 208). Although statistical methods of diagnosis and sensitivity tests exist, these data problems are best addressed by moving from a complete reliance on gathering data from news sources to incorporating supplementary sources.

Several data sets that record protest events exist, varying in temporal and geographical coverage, as well as in source material. One of the reasons to why I chose ACLED over other protest datasets is the comprehensiveness of the source material used to construct the dataset. Other data on protest, such as the Nonviolent and Violent Campaigns and Outcomes (NAVCO 3.0), Integrated Crisis Early Warning System (ICEWS), the Social Conflict Analysis Dataset (SCAD) and the Mass Mobilization in Autocracies Database (MMAD), rely primarily information reported in the media. Thus, the potential issue of reporting bias is evident. ACLED too draws on daily local, regional, national, and continental media reports in over 20 languages. However, these are explicitly supplemented by NGO and IGO reports, social media platforms, and partnerships with local informants and observatories in under-reported cases (*Armed Conflict & Event Data Project (ACLED) Codebook, 2019* 2019, 33).

A second reliability issue is addressed in an earlier article by Weidmann (2014). By comparing media- and military-based event data, the author shows that measurement errors regarding location and casualty numbers of events increase the more remote the location of the event is (Weidmann 2014, 1140–4). The author further argues that using geographical information on event location is generally reliable on the district level, but not below (Weidmann 2014, 1143). In order to alleviate this concern of spatial inaccuracy, I follow the author's advice to move away from specific coordinates to a higher level of aggregation. I do so by using the PRIO-GRID framework, which I turn to next.

4.2 Unit of analysis: PRIO-GRID

Each event in ACLED is assigned latitude and longitude coordinates. I have chosen to overlay this with the widely used PRIO-GRID spatial framework (Tollefsen, Strand, and Buhaug 2012). PRIO-GRID is a consistent grid structure that divides the world into 0.5×0.5 degree cells. This corresponds to approximately (50×50) kilometers^2, or an area of 250 square-kilometers, on the equator (Tollefsen, Strand, and Buhaug 2012, 367). Using this type of structural framework provides a solution to the fact that ACLED events are not assigned identifiers that connect observations, which is necessarry in order to assess change.

Furthermore, it addresses the concerns of spatial inaccuracy of event data in remote areas (Weidmann 2014). Nonetheless, the resolution is sufficiently fine so that even small countries consist of multiple cells, allowing for the examination of subnational patterns. The grid cells are inherently apolitical as well as identical across countries (as opposed to district-level administrative units). Thus, the unit of observation is "completely exogenous to the feature of interest" (Tollefsen, Strand, and Buhaug 2012, 365).

In other words, all events recorded to take place within the same grid cell are grouped together. ACLED also codes the geographical precision level. In order to avoid the risk of misclassifying event location, I have filtered out events that have the lowest spatial precision code. The sources of these events typically mention a larger region as the protest site (*Armed Conflict & Event Data Project (ACLED) Codebook, 2019*

2019). This is the case for a very small number of observations. Events that are reported to take place within a small part of a region (spatial precision code = 2) are kept, as these are more likely to fall within the "correct" (50×50) km^2 grid cell area.

By grouping observations in space and time, the analysis is not centered around specific protest movements or campaigns and their political aims. Rather, I focus on characteristics of the overall protest activity on the subnational level as well as government responses to this, offering a novel conceptualization of protest. It represents a more objective measure of protest activity for two reasons. Firstly, I do not ask the question of "what constitutes a campaign or movement?". Although there are conventional answers to this question, a subjective evaluation is always present. Secondly, I do not select the cases to be observed based on the answer to this question. Rather, the sample is based upon objective evaluations of whether events occur at the same place at the same time.

Hence, the unit of analysis is not the protest event, but rather the grid-cell. Although the data would allow to observe the grid cell on a day-to-day basis, doing so raises concerns about whether there is sufficient variation in the dependent variable (S. R. Bell and Murdie 2016). In order to avoid excessive temporal aggregation, I follow Rasler (1996) and S. R. Bell and Murdie (2016) and choose the week as my temporal unit of observation. As I focus on the what happens in between onset and outcome in this thesis, this allows for a fine-grained analysis of the sequence of events.

4.3 Protest cycles

The weekly grid-cell structure was created by recoding the date in which the event took place from the day to the week. To illustrate, in Cape Town during the week of October 3, 2016, at the height of the 2016 #FMF protests, eleven unique demonstrations at different points in time and space are coded in ACLED. These locations fall into two different grid cells, which illustrates the apolitical nature of PRIO-GRID. In one of the grid cells, covering the University of Cape Town (UCT), seven events are recorded; two of which were violent demonstrations and five nonviolent demonstrations. These data points are then grouped together, forming one observational unit.

In order to measure change over time, I have grouped protest weeks into protest cycles. Defining the sample required a considerable amount of data tidying. In the final dataset, individual grid cells are observed in all weeks where there is ongoing protest activity, given that the observation does not represent a single protest week. It is not possible to measure *change* on the chosen level of aggregation if the unit of analysis is measured only once. Weeks without protest are filtered out of the final dataset, because they have missing values on the dependent variable. Because no change is observed in the start week, this week is also removed when estimating the models. Somewhat similar to a campaign, the protest spells have discernable start and end points of protest activity within the grid cell, as well as distinguishable events from start to finish (Chenoweth and Stephan 2011, 14). As I am interested in modeling violent escalation, rather than the choice of protesting (violently), protest cycles that are totally violent in the first week of protest are excluded from the sample, as are weeks with no registered events.

There is no straightforward solution as how to define the start and end point of the cycle. One possibility is to define protest cycles as strictly consecutive weeks of reported activity. Here, a cycle is an uninterrupted series of protest weeks. If we observe protest in grid-cell week$_s$, but no protest in week$_{s-1}$, week$_s$ is defined as the start week of the spell, given that there is also protest in week$_{s+1}$. Similarly, given that there is observed protest in week$_{e-1}$ and week$_e$, but not in week$_{e+1}$, week$_e$ is defined as the end week. Grid-cell weeks that fall within the start and end week are thus part of the same cycle. A single week of no reported activity marks the end of the cycle.

One clear limitation with this strict definition is that many of the protest cycles were separated by only one week. Because protests are observed within a relatively small area, it is difficult to treat these observations as independent from one another. While this is of little importance as I am measuring change from week to week, the duration of protest is a possible confounding factor that could potentially be biased.

The challenge then becomes to define an accurate number of weeks of non-activity that is required to define start and end weeks of protest cycles. Instead of the one-week threshold, I chose a two-week threshold, where the start week is week$_s$ if protest is not observed in neither week$_{s-2}$ nor week$_{s-1}$ but is observed in week$_{s+1}$ and/or week$_{s+2}$. Likewise, end weeks$_e$ are given by no observations of protest in neither week$_{e+1}$ nor week$_{e+2}$. Hence, two consecutive weeks of no protest marks either the start or end of a spell. Using this definition, the dataset contains 5779 distinct cycles that occur in 979 grid cells, comprising a total of 20232 protest weeks.

4.3.1 Country coding

One problematic aspect of aggregating protest events to the grid-cell level is that some grid cells cover the territory of multiple states, which is frequent in border areas. In PRIO-GRID, each cell is allocated to one state. In multiple-country cells, they are assigned to the country that covers the largest proportion of the area of the cell (Tollefsen, Strand, and Buhaug 2012). However, some protest cycles take place in the country that happens to cover a minority share of the area of the grid cell. I therefore relied on ACLED's coding of the country in which the protest events occurred when incorporating the country-level control variables into the dataset.

While this is not of much concern if protest events only occur on one side of the border at a given time, a problem arises when protest events take place in the same grid cell within the same period of time, but on different sides of the border. The most illuminating example of this problem concerns a particular grid cell that covers a part of the border between the Democratic Republic of Congo (DRC) and the Republic of Congo (PRIO-GRID ID = 123511). According to PRIO-GRID's majority rule, this grid cell "belongs" to Congo. The capitals of DRC and the Republic of Congo, Kinshasa and Brazzaville, are separated only by the Congo River, and happen to fall within the same grid cell. Hence, protest events taking place both in Kinshasa and Brazzaville during the same time period are automatically considered as events of the same cycle. Although the distance between Kinshasa and Brazzaville is short, assuming that protesters in one capital would be affected by government repression of a protest perpetrated by a government in a different country would be speculative.

In order to address this problem, I manually inspected each of the 13 grid cells that included protest events taking place in multiple countries at the same time. In the simplest cases, only one event occurred in a different country from the others in the cycle. Here, the single event was removed in order to ensure consistency. Similarly, if there was only one protest week during the cycle that had events in both countries, I filtered out the events from the country that did not have protest during the other weeks of the cycle. In cases where the cycle consisted of two single protest weeks, each of which occurred in different countries, the entire cycle was removed from the dataset. The most complicated cycles involved several protest weeks with events in two countries. This was the case for two cycles in the sample. One cycle included six protest weeks in DRC and four weeks in the Republic of Congo. Here, the four protest weeks in the Republic of Congo were removed as the protest began 100 percent violently. The second cycle took place on the border between DRC and Burundi. In this case, the protest weeks in DRC occurred with a two-week gap and would hence not constitute a new cycle. These protest weeks were therefore removed.

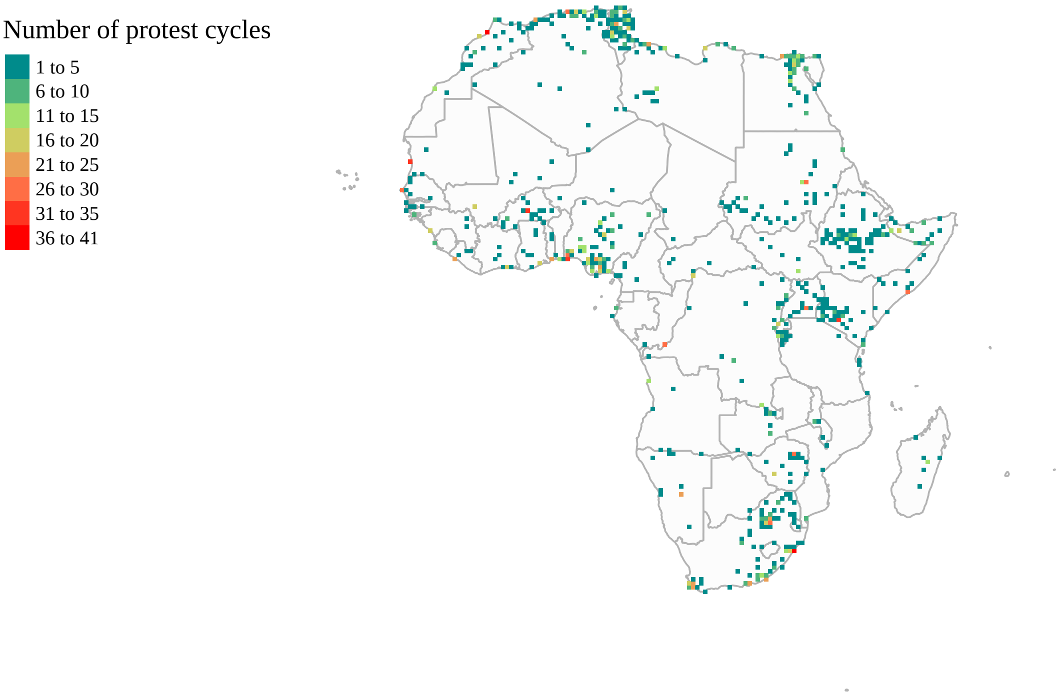

Figure 5: Protest cycles in Africa, 2010-2018 (ACLED)

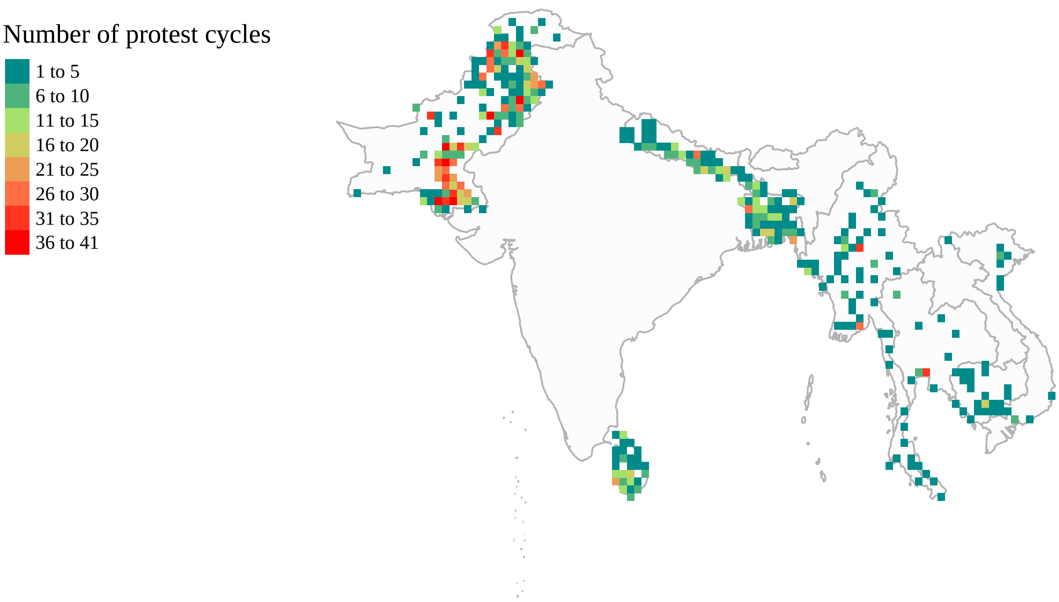

Figure 6: Protest cycles in Asia, 2010-2018 (ACLED)

4.4 Dependent variable: Change in proportion of violent events

To operationalize the dependent variable, I first divided the number of violent protest events by the total number of events, thus obtaining the proportion of events that are violent in a given grid-cell week. This is similar to the method used by Cunningham and Beaulieu (2010) to capture strategic choices made by dissident groups. Where others have counted the number of (violent) protest events conducted by the same campaign (Pinckney 2016), the onset of a violent phase by a campaign (Shellman, Levey, and Young 2013), or a binary measure of escalation of single events (Gustafson 2019), my dependent variable allows me to assess the relative change in violent activity on a standardized scale that is comparable across units, ranges from 0, totally nonviolent, to 1, totally violent (Bond et al. 1997). Secondly, I extend this operationalization by looking at change from week$_{t-1}$ to week$_t$, rather than the proportion in week$_t$.

I now turn to the definitions of nonviolent and violent protest, drawn from the ACLED codebook [@-ACLED2019]. *Nonviolent protest* is defined as any protest event where the protesters do not use violent or disruptive tactics. This includes both one-sided and two-sided peaceful protests, as well as the sub-event types "protest with intervention" and "excessive force against protesters". In the latter two instances, the protesters are not reported to have used violent means and the event is therefore counted as a nonviolent protest event. Two-sided protests are e.g. rival protests for and against a given cause. These events are counted as one protest event in the data. One-sided protests are the most general form of peaceful protest in which there is no interaction between the protesters and a second actor.

Violent protests are drawn from from ACLED's category of "violent demonstration". The actors engaged in violent demonstrations are defined as rioters. Violent demonstrations are characterized by their use of violence, defined as disruptive acts. The threshold of violent behavior in ACLED is low, in line with the broad definition of violence above. The disruptive acts not only include physical violence targeted at persons, but also acts of vandalism, road-blocking, barricading, tire-burning, and other acts of violence targeting property or businesses. This also includes events of violent protest that engage a second actor, such as police forces.

The *total protest activity* is thus the sum of all nonviolent and violent protest events occurring within the same grid-cell during the same week.

The dependent variable is therefore operationalized as the weekly change in the proportion of violent events. That is, $change = PV_t - PV_{t-1}$, where PV_t is the proportion of violent events in week$_t$ and PV_{t-1} is the proportion of violent events in the same grid cell one week earlier. This means that the scale of the dependent variable ranges from -1 to 1, where negative values represent a decrease in violent protests, relative to the total number of protest events, and positive values indicate that violent protest events represent an increasing proportion of the total protest activity. The dependent variable therefore captures tactical shifts rather than escalation in terms of an increse in the number of events. As robustness checks, I also estimate models using various specifications of the dependent variable, e.g. coding all decreases as zero and only modeling escalation.

4.5 Independent variable: Government repression

To test the substitution hypotheses, a measure of the occurrence and frequency of government repression is required. Government repression is operationalized using ACLED's category of protest events where an interaction between the protesters and state forces is coded. This typically includes police and military forces that are recognized agents of government.

Repression of nonviolent protest is operationalized using ACLED's two sub-types of nonviolent protest, "protest with intervention" or "excessive force against protesters". In the former case, state forces use non-lethal techniques of intervention to suppress the protest, such as crowd dispersal. Furthermore, there are no reports of serious injuries or casualties. While not lethal, this falls into the conventional definition of repression as crowd dispersal is a coercive measure that increases the protester's costs of contention (Davenport 2007). The category "excessive force against protesters" is used where the nonviolent protest is targeted with lethal weapons or the security forces use other forms of violent techniques that have the potential to inflict serious or lethal harm on the protesters. To capture the possible differential effects of repression, I include both types of repression of nonviolent protest in the model, thus explicitly separating mild from harsh repression of nonviolent protesters (Shellman, Levey, and Young 2013). Thus, in the following, when I refer mild and harsh repression of nonviolent protest, these concepts refer to events in ACLED that are coded as intervention and excessive force, respectively. A mere 6.4 percent of the total nonviolent protest events in the data were repressed either mildly or harshly.

Government repression of violent protest is oprationalized using ACLED's coding of any interactions between violent protesters and police or military forces. Where the violent protest is coded to have involved an interaction with these state-controlled bodies, this is coded 1 and 0 otherwise. Ideally, I would distinguish between different forms of repression akin to the two categories above. Unfortunately, ACLED lacks the nuance between non-lethal and (potentially) lethal repressive acts towards violent protests. In the sample, 46.4 percent of the individual violent protest events were repressed by police or military forces. Unsurprisingly, repression of violent protests is far more common than repression of nonviolent protest.

Repression of nonviolent or violent protests may be perpetrated by other actors than police or military forces. The other types of actors in ACLED include rebel groups, political or ethnic militias, and private security forces. In order to capture the effect of government-led repression on protest tactics, interactions of these types are excluded from the variables. Although I acknowledge that there may be an independent effect of the use of violence against protesters carried out by other actors, the number of events where this occurs is too small to accurately to discern this.

In order to test the effect of the frequency and severity of government repression in the short term, I therefore constructed three variables that measure the level of repression in the previous grid-cell week. Like the dependent variable, the repression variables are relative measures of repression to facilitate comparison. The frequency of mild repression of nonviolent protest; harsh repression of nonviolent protest; and repression of violent protest, respectively, are divided by the *total number of protest events* to obtain the proportion of repressed events. All three variables are lagged one week. This is done to avoid problems of endogeneity within the model itself and retaining the ability to assess the effect of repressive actions in the very short term.

4.6 Independent variable: Level of organization

In ACLED, protesters and rioters are coded using generic categories, i.e. "Protesters (country)" and "Rioters (country)". If any, ACLED also identifies named groups that are affiliated with the protesters in a separate column, 'associated actor'. This means that if a group is known to lead an event, the name of that group is recorded.

I have used this information to operationalize the variable organized protest to test H_2. The variable is constructed as follows. Observations where a named group is known to be affiliated with or leading a nonviolent protest event, are coded 1. This refers to protests where there is some level of organization behind the protest event. Conversely, if no associated groups are identified for a protest event, i.e. the associated actor column is empty, the variable is coded 0. In other words, protests are considered to be disorganized where no named actor is coded to have organized the event. Aggregated to the grid-cell week, protest organization is operationalized as the sum of the nonviolent protest events that identify an active group divided by the total sum of protest events.

The coding of associated actors in ACLED is somewhat inconsistent when more than one actor is involved. Sometimes the protesters or rioters are coded as actor 1, and other times as actor 2. A second source of unreliability concerns the handful of observations in the original data where the group associated with protesters were identified as "civilians". From this, one cannot readily draw the conclusion that there is any form of organization behind the event. This contrasts cases where e.g. "teachers" are identified as the associated actor. While "teachers" do not represent a named organization or group, it is more likely that the event is, to some extent, organized.

A substantial amount of data tidying was therefore necessary. I examined only the cases in which either actor 1 or actor 2 were categorized as "Protesters (country)". From this, I coded the dummy variable according to whether any associated and named group, apart from "civilians"[6], was identified, that was used to construct the proportion variable.

4.7 Exemplifying the dependent and independent variables

To exemplify the operationalizations above, I consider one specific grid cell that covers Cape Town, South Africa, in two consecutive weeks during the fall of 2016. In the first week, the week of October 3, a total of seven protest events are coded. Of these, five were nonviolent while two are coded as violent protest. Thus, the total proportion of violent protests is $\frac{2}{7} \approx 0.286$. In this example, this is the proportion of violent protests in week$_{t-1}$.

Of the five nonviolent protests, one is coded as "protest with intervention": on October 3, 2016, students at the UCT protested by blocking the building's main entrances, which led some of them to be arrested. The following day, at the same location, police forces fired tear gas and stun grenades on rioting students who had set fire to university buildings. The former example is coded as mild repression of nonviolent protest, because of the arrest of protesters who did not use violent means. The protest on October 4 is coded as repression of violent protest, as there was significant interaction between the police and the protesters who did intentionally damage property, in line with the broad definition of violence.

Turning to the second week, the week of October 10, one nonviolent protest and five violent protests are coded. As the nonviolent protest is not recorded to have been organized by a named group, the proportion of nonviolent protests that are organized equals zero. This week, the proportion of violent protests is $\frac{5}{6} \approx 0.833$. For this specfic grid-cell-week, the dependent variable is therefore $PV_t - PV_{t-1} = 0.833 - 0.286 \approx 0.547$; a sharp increase from the previous week.

[6] Civilians are defined as "those who do not actively choose to be involved in an event", and are therefore excluded (*Armed Conflict & Event Data Project (ACLED) Codebook, 2019* 2019, 18)

4.8 Control variables

In this section, I operationalize the control variables used in the analysis and describe the data sources. While I in this thesis primarily focus on short-term changes that may lead to violent escalation, I recognize that protests occur within a political environment and context. I therefore include a number of control variables that have been found to increase the risk of political violence. Most of the control variables are therefore measured on the country level. Rather than directly influencing the risk of violent protest escalation, the control variables may rather exert an indirect effect by influencing the relative capacity of governments to repress or the protester's incentives to protest in the first place.

4.8.1 Structural variables

I begin by presenting the control variables that measure different structural attributes of the polity in which the protests occur. Thus, these variables are relatively static.

GDP per capita

Gross Domestic Product (GDP) per capita acts both as a proxy for state capacity, which may influence its ability to repress, and as a proxy for income-related grievances. The data for GDP per capita is collected from the United Nations Statistics Division ("Per Capita GDP at constant 2015 prices in US Dollars" 2020). To my knowledge, no spatially *and* temporally disaggregated data on income exists. The trade-off is between standard country-year data and subnational-level data that is only measured or modeled in one year (e.g. 2015). I therefore chose to rely country-level yearly data, assuming high levels of correlation within the country, although this is by no means a perfect measure. GDP per capita is measured in constant 2015 U.S. Dollars. While there are no assumptions placed on the distribution of independent variables in linear regression, heavily skewed variables, which is the case for this variable, may lead to heteroscedastic and non-normally distributed residuals (Christophersen 2013). Hence, the variable is log-transformed.

Urban-rural

According to Gustafson (2019), the risk of violent protest escalation is higher in rural areas because of the increased risk of protest collapse in less-populated areas, which is controlled for using a binary variable. While this may be true, protests in urban areas may pose a higher threat to the government, which may in turn lead to repression. As repression raises the costs of protest, protests in urban locations may thus indirectly lead to a higher risk of violent escalation. Moreover, contrary to the evidence presented by Gustafson (2019), protests in more populated urban areas may be able to attract higher number of participants, which may lead to fractionalization and a higher overall risk of violence.

To control for urban-rural divisions, I use the Malaria Atlas Project's (MAP) Accessibility to Cities raster data (Weiss et al. 2018). The MAP data maps travel time to high-density urban centers[7] in a 1×1 kilometer raster. This raw data is thus substantially more fine-grained than the 50×50 kilometer PRIO-GRID framework. To overlay the travel time data to the PRIO-GRID cells, I used an aggregation function that computes the average travel time within the cell. The MAP data is available only as a snapshot for 2015 and is therefore static across grid cells. In urban areas, the average travel time is considerably shorter than in rural areas and comparatively less developed areas. Inspecting the data, most protests occur in relatively urban areas. Because the distribution is significantly right-skewed, I log-transformed the travel time variable.

Regime characteristics

I include two central characteristics of the regime under which the protest occurs. An increased willingness for violence may stem from restricted opportunities for political participation and advancing political claims by nonviolent means. To control for regime type, I included the liberal democracy index from the Varieties of Democracy (V-DEM) dataset, version 9 (Coppedge et al. 2019). The index is measured on the country-year level with a scale ranging from 0 (low) to 1 (high). I chose to measure

[7]Urban centres are defined as "a contiguous area with 1,500 or more inhabitants per square kilometer or a majority of built-up land cover coincident with a population centre of at least 50,000 inhabitants" (Weiss et al. 2018, 333).

regime type using the concept of liberal democracy because it not only captures the extent to which electoral democracy is achieved, but also incorporates the limits placed upon the government and the degree to which individual and minority rights are protected. Crucially, these include civil liberties such as the right to speech and to assemble. Both in the civil war and the protest literature, it has been shown that semi-democracies are more likely to see higher levels of both protest and violence compared to strict autocracies and full democracies (Hegre et al. 2001). I therefore included both a linear and a squared term of the democracy measure.

From the above hypotheses, I expect that, all else equal, government repression is a significant predictor of whether nonviolent protests will turn violent. However, there is significant variations in governments' willingness to use potentially lethal repressive acts on their citizens. In addition to regime type, I control for regime strength for two reasons. Politically stable leaders may be less likely to repress, and protest is less likely to pose a threat to strong regimes. I used the Rulers, Elections, and Irregular Governance (REIGN) dataset to operationalize regime strength (C. Bell 2016). Strong leaders are assumed to have a long tenure, measured as the number of months the current leader has been in power. Thus, the count variable resets in months where a leadership or reelection occurred. This control variable is thus on the country-month level. I also included a dummy variable to control for whether the current leader's rule is to some extent legitimized through an election[8]. Both the tenure and the elected variables are lagged one month to avoid problems of endogeneity.

Unemployment rate

Gustafson (2019) finds that high unemployment rates increase the risk of violent escalation[9]. It is somewhat problematic that the author uses a country-year level variable to address his hypothesis stating that as a country's unemployment rate *increases*, the risk of violent escalation increases. However, to my knowledge, no

[8]Note: This does not necessarily mean that the election has been competitive or democratic.

[9]The author's second key variable, food-price increases, is excluded as the monthly ILO consumer price data is missing for a substantial number of the country-months in my dataset. The missing data cannot be assumed to be missing completely at random. Rather, it is systematically missing for some countries and regions as a whole.

temporally disaggregated data on unemployment rates is available that covers all countries in my data. To control for this effect, I therefore use the modeled International Labor Organization (ILO) estimates, retrieved from the World Bank ("Unemployment, total (% of total labor force) (modeled ILO estimate)" 2020). By using the modeled estimate, self-reporting biases and missing data problems are alleviated (Gustafson 2019).

4.8.2 Protest variables

I now turn to the control variables that are measured on the same grid-cell-week level as the protests. These covariates tap into different characteristics of the protest and are assumed to also influence the risk of increasing violence.

Proportion of violent protest in week$_{t-1}$

The most parsimoneous model that I report in the analysis below contains only the level of violent protest within the grid cell in week$_{t-1}$. First, it is thought that the level of violence in week$_{t-1}$ may have an independent effect on dependent variable which is measured one week later. Second, repression of nonviolent protest in very violent environments may have a different effect than repression of nonviolent protest where there are no violent protests occurring. Third, repression of either type of protest may be more likely in very violent situations. Controlling for the proportion of events that are violent, lagged one week, thus allows for assessing the individual effect of repression.

Duration

The duration of the protest activity in the grid cells in my sample varies considerably. While only 34 of the 5779 protest cycles in my sample lasted 52 weeks or more, and 32 percent of protest last only two weeks, some grid cells have seemingly been entrenched in protest during the past decade.

In spontaneous, or disorganized, protests, there is often a significant uncertainty to how long the protest activity will last. This is thought to increase the risk of protesters turning to violent means in fear of collective action collapse (Gustafson

2019). In other words, the risk of violent escalation may be highest in the beginning of protest actions. However, it is also conceivable that the longer the duration of (unsuccessful) protest, the more impatient protesters grow, which in turn may also lead them to turn to violence.

To control for these different effects of protest duration, I constructed a variable, *duration*, that counts the number of weeks since the first protest *in that cycle* started for each grid-cell week. This variable therefore follows my definition of protest cycles where an interruption in protest activity that exceeds one week marks the end of a cycle. The average duration of a protest cycle is 5.34 weeks. Because the variable is measured cumulatively and most cycles last for a relatively short period of time, its distribution is right-skewed. The duration variable is therefore log-transformed.

Spatial diffusion

Rasler (1996) argues that protest spreads by spatial diffusion, increasing the total level of protest activity. The within-country spread of violent protest adds pressure on the government and is thought to increase their incentives to repress dissent. Furthermore, the total level of protest activity within the country increases the likelihood of violent escalation in itself.

To account for this this, I constructed two variables. Each grid cell is adjacent to eight other grid cells. In other words, each cell has eight first-order neighbors. Following Weidmann and Ward (2010), I only use first-order neighbors because diffusion effects from higher-order neighbors are hypothesized to go through their neighbors, so that the effect will be captured by the first-order neighbor variable.

To create the variables, I first identified the neighbors using the *pgneighbors* function from the PRIO-GRID version 3 GitHub repository (Vestby et al. (2020)). After joining this with the ACLED data, I was able to identify whether there was ongoing violent protest in any of the grid cell's eight neighbors in any given week. The first variable, *protest in neighboring grid cells*, measures the extent of violent protest in each unit's immediate neighborhood in any given week. This is operationalized as the count of the number of grid cells where violent protest occurs in week$_t$.

The second variable, *diffusion*, measures the spread of violent protest. This is operationalized as difference of the number of adjacent grid cells with ongoing violent protest between week$_{t-1}$ and week$_t$. This takes a positive value if there is more violent protest in the neighborhood compared to the previous week, a negative value if less is observed, and zero if there is no change.

Lagged dependent variable

Governments are more prone to repress violent protest rather than nonviolent protest (S. R. Bell and Murdie 2016; Cunningham and Beaulieu 2010). Repression is therefore more likely to occur when and where there is an increase in the level of violent protest. Similarly, where we observe a decrease in the proportion of violent protest, government repression of any kind is less likely. Therefore, in some of the models, I include the dependent variable, change in proportion of violent events, lagged one week as a control variable.

Apart from the substantial reasoning to control for changes in relative violence, this is a well-known method to counteract serial correlation (Christophersen 2013). Serial, or auto-, correlation occurs when the error term, e_t, is strongly correlated with e_{t-1}, which is likely in my data as it has an unbalanced panel structure (Pevehouse and Brozek 2008).

Table 2 shows the distribution of variables.

Table 2: Descriptive statistics

Statistic	N	Min	Max	Mean	St. Dev.
Change in proportion of violent events	20,232	-1	1	0.058	0.339
Mild repression (NVP) in week$_{t-1}$*	20,232	0	1	0.038	0.167
Harsh repression (NVP) in week$_{t-1}$*	20,232	0	1	0.012	0.094
Repression (VP) in week$_{t-1}$*	20,232	0	1	0.036	0.153
Organized protest*	20,232	0	1	0.481	0.433
Duration (log)	20,232	0	6	1.733	1.428
Protest in neighboring grid cells (lagged)	20,232	0	6	0.639	0.925
Spatial diffusion (lagged)	20,232	-3	4	-0.001	0.321
GDP per capita (log)	20,232	4.650	9.980	7.432	0.717
Average travel time (log)	20,232	-1.368	7.016	3.166	1.201
Liberal democracy	20,232	0.030	0.693	0.313	0.153
Liberal democracy^2	20,232	0.001	0.480	0.122	0.115
Monthly leader tenure (lagged)	20,232	1	498	55.074	74.984
Elected (lagged)	20,232	0	1	0.557	0.497
Unemployment rate	20,232	0.273	27.327	6.826	7.360

* As proportion of total protest events

4.9 Research design: Multi-model approach

I now turn to present my research design. Choosing the appropriate regression model for the dataset I have compiled is not completely straightforward. I have chosen to estimate a number of Ordinary Least Squares (OLS) regression models with various specifications. The linear model is the most widely used modeling strategy, and with a metric dependent variable of this kind, the linear model is found to be a robust estimator in a wide variety of settings (Angrist and Pischke 2009).

The estimated additive linear model is given by

$$\tilde{Y} = \beta_0 + \hat{\beta}_1 mildrepression_{NVP} + \hat{\beta}_2 harshrepression_{NVP} + \hat{\beta}_3 repression_{VP} + \hat{\beta}_4 organization + \hat{\beta}_k X_k + e_i,$$

where $\tilde{Y}$ denotes the predicted value of the outcome variable, given the predictor variables, X_k, and the associated estimated coefficients, $\hat{\beta}_k$. The individual error term, e_i is given by the difference between predicted and observed values on the dependent variable. By minimizing the sum of squared errors, the model estimates the coefficients that produce the best fitting line between the independent and dependent variable(s) (Stock and Watson 2007, 118). As such, the multivariate OLS model works well to assess the strength and form of the controlled association between covariates and the metric dependent variable. Simple regression models are generally preferable to more complex and nonparametric estimators (Angrist and Pischke 2009). OLS approximates the population average of the dependent variable, given the independent variable(s), along with the uncertainty of the results, even in instances where the relationship in the data is nonlinear (Angrist and Pischke 2009, 2010). Hence, the patterns in the data are approximated by the linear model. The standard errors that reflect the uncertainty of the results are elaborated upon below.

The dependent variable, change in the proportion of violent protest, is measured on a scale from -1 to 1. Negative values indicate that there has been a decreasing proportion of violent events, while positive values indicate an increase, relative to the total protest activity. Figure 9 shows the distribution of the dependent variable, which is near-normally distributed. From the figure, it is evident that most grid-cell weeks have the same proportion of violent event as the previous week. In these cases, there is no change and the dependent variable equals zero. As a robustness check, I also report the results of a model where only increases of the dependent variable is estimated, i.e. where $Y_i > 0$. Here, the scale ranges from 0 to 1. Positive values of the dependent variable indicate that there has been an increase in the proportion of violent protests.

In order to assess the robustness of the results and whether the results are products of the constraints imposed by the linear model, I also estimate logistic regression models and various specifications of the OLS model (reported in Appendix). This approach allows me to test whether the results hold across different model specifications, while acknowledging that there may not be a single model that perfectly handles all the potential issues or accurately models all data (Angrist and Pischke 2010). Thus, instead of relying on one specific model, I have chosen to start out with a set of parsimoneous regression models and add control variables in a step-wise manner

(Christophersen 2013, 65). The interaction terms used to assess H_3 are more closely inspected in chapter 5.2. Following the empirical analysis, I discuss the statistical validity of the reported results.

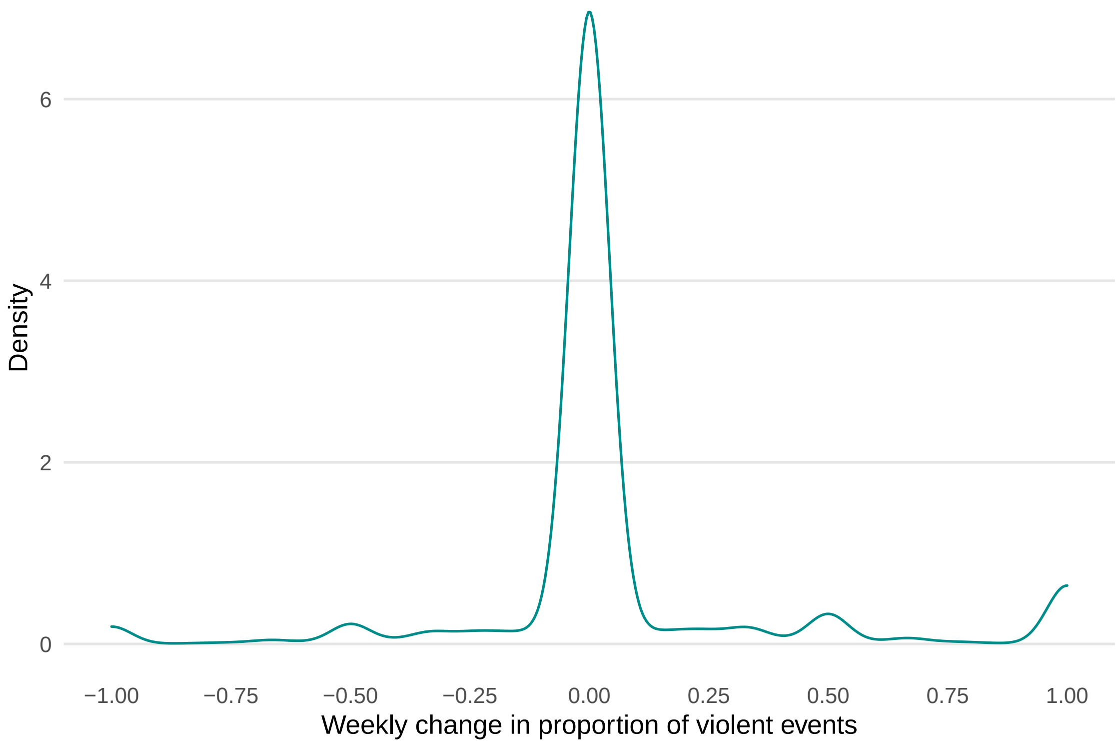

Figure 7: Distribution of the dependent variable

4.10 A note on standard errors

Along with the coefficient estimates, precise standard errors is critical for testing hypotheses, determining the statistical significance of variables, and drawing valid statistical inferences (Cameron and Miller 2015). The standard error that accompanies a regression coefficient is a measure of how much the estimated coefficient varies, and is calculated using the squared sum of residuals (Christophersen 2013). Poorly fitted models produce large residuals, which in turn lead to large standard errors and insignificant results. Because of the large sample I have compiled for this analysis, the standard errors are likely to be small with precisely estimated coefficients. However, counteracting this is that there is relatively little variance (squared standard deviation) in the independent variables of interest. Furthermore, there are issues with the data that can lead to artificially small standard errors. In this section, I highlight some of these problems that can lead to misspecified standard errors, along with the solutions that are implemented to alleviate these concerns.

The structure of the dataset is longitudinal, in the sense that grid cells are observed in weeks where there is ongoing protest. 17 percent of the grid cells in the dataset only have one protest cycle, while the remaining 83 percent are observed in multiple protest cycles. As the duration of the cycles vary and grid cells are only observed in weeks of protest, the data can best be described as an unbalanced panel. The assumption that the observations, i.e. the observed grid-cell weeks, are independent and identically distributed (iid.) units, is therefore difficult to defend. Grid-cell weeks are nested in protest cycles, that are also nested in grid cells, that again are within countries. In addition, some of the control variables are measured on the country level.

Autocorrelation, one of the base assumptions for OLS, is a well-known problem in panel data (Worrall 2010). This occurs when there is signficant correlation between residuals over time, indicating that there is temporal dependency in the data. Although my primary focus is on the short-term impact of covariates, there may be unobserved and/or largely static characteristics of the grid cells that influence the likelihood of violent escalation of protest. Some structural control variables are included in the full model, yet the small variance of these within the units is problematic. One way to assess the presence of autocorrelation is to estimate a linear model with the residuals from the original model is used as the dependent variable and a lagged version of the residual as an independent variable, alongside the original covariates (Worrall 2010, 188). A significant coefficient of the lagged residual indicates serial correlation, which I found to be the case in my data. While autocorrelation does not influence the estimates of the coefficients, it can lead to misspecified standard errors.

All the regression tables below are estimated using clustered standard errors, which is one way to account for autocorrelation in panel data. Because of the clusters in the data, the number of observations in the dataset may be effectively smaller than the 20232 number of observed grid-cell weeks (Johnston 2008). Protests occur, and reoccur, within a specific political environment. If one does not take this into account, standard errors can be underestimated, which consequently leads to narrow confidence intervals, large t-statistics, and low p-values, all of which increases the risk of Type-I error (Cameron and Miller 2015; Johnston 2008). Type-I error occurs when a true null hypothesis is rejected based on falsely signficant estimates. In order to choose the appropriate clustering unit, I followed the advice of Cameron and Miller (2015) and compared the standard errors clustered at different levels; the grid cell and the country. Because the standard errors increased substantially when clustering on the country-level compared to on the grid cell, I chose the country level to avoid underestimating the standard errors. Within-country autocorrelation occurs when residuals correlate over time within the country and across grid-cells, but are assumed to be uncorrelated between countries (Cameron and Miller 2015).

A second way to account for the autocorrelation in the data is to estimate an autoregressive model (Christophersen 2013). Here, the dependent variable is lagged by one or more temporal units and included as a covariate in the model (Worrall 2010). The first-order autoregressive model (AR(1)) includes the dependent variable lagged one week, Y_{t-1}, as a control variable. In addition to account for autocorrelation, the lagged dependent variable also controls for omitted variables, reducing the risk of omitted variable bias (Christophersen 2013). For each observation, the lagged dependent variable controls for all omitted variables up until week$_{t-1}$, but not from week$_{t-1}$ to week$_t$. Hence, it is important to note that this models only short-term changes. This is of little concern to my model, precisely because of the hypothesized "immediate" impact of repression and organization.

However, one drawback with the autoregressive model is a loss of observations. This is because the first group-wise dependent variable cannot be lagged. Because the dependent variable, *change*, is operationalized as the difference between week$_t$ and week$_{t-1}$, the very first observation within each protest cycle is excluded from the start, as explained above. However, introducing the lagged dependent variable as a way to account for autocorrelation leads to a removal of the first observation, after this filtering. In other words, there is a substantial reduction of observations in this model (N = 14 453 versus N = 20 232). This further reduces the number of unique grid cells that are observed from 979 to 547, and the number of countries that are observed from 54 to 51. In addition, only protest cycles that last longer than three weeks are included - and only observed from the third week onwards. Despite this caveat, the sample remains large, covering a wide variety of polities.

A third way to account for autocorrelation, which in essence is omitted variables that are static across time or groups, is to incorporate fixed effects (Angrist and Pischke 2009). This involves estimating a dummy variable for the grouping unit in the model. Spatially, it is not straightforward whether the fixed effects should be estimated on the grid-cell level or the country level (or any other administrative level). Temporally, monthly and weekly fixed effects are as plausible as yearly fixed effects. Below, I report the results from the grid-cell fixed effects model, as the data is organized as a panel structure with the grid cell as the observational unit, in addition to a country-year fixed effects model (other fixed effects models are included in Appendix, with similar results). An F test showed that the grid-cell fixed effects model is preferable over the standard OLS model, as there is significant individual heterogeneity in the data (see Appendix).

A caveat with the fixed effects model is that it only allows for within-group comparison (Mummolo and Peterson 2018). The estimated coefficients answer to the change in Y by a within-unit shift of the associated predictor. If the omitted variables that are modeled through the use of fixed effects vary across the units, this is not possible to assess (Christophersen 2013). A second issue relates to the loss in degrees of freedom when estimating unit-specific dummy variables[10]. However, as opposed to the autoregressive model, one avoids the loss of observations. The choice between fixed effects and autoregressive models is not clear-cut. I therefore follow the advice of Angrist and Pischke (2009) and estimate both models in order to assess the robustness of the results. This is done in addition to the standard OLS models.

[10]This trade-off could be alleviated by estimating a random effects model, but the significant Hausman test showed that the fixed effects model is preferable over random effects (see Appendix).

5 Empirical analysis and discussion

In this chapter, I present the empirical analysis of the data. The aim of the empirical analysis is to test the hypotheses regarding associations between organization, repression and violent escalation and de-escalation of nonviolent protest. First, I inspect the baseline models, which are several specifications of the linear model estimated without control variables, including the first-order autoregressive model. I proceed with incorporating the controls for several structural attributes of the political environment. In section 5.2, I include the interaction terms in order to assess H_3 and provide a thorough inspection of the interaction effects. Here, I show that the association between repression and violent escalation is conditional upon the level of organization. In section 5.3, I assess the strength of the association. Finally, I summarize the results by discussing of the validity of the results in terms of drawing statistical inferences, and whether the results can be interpreted in causal terms.

Table 3: Baseline models

	Dependent variable:								
	Change in prop. of violent events in $week_t$								
	Baseline						AR(1)	FE	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Constant	0.105^{**}	0.101^{**}	0.105^{**}	0.105^{**}	0.235^{***}	0.231^{***}	0.197^{**}	-0.000	0.199^{***}
	(0.036)	(0.036)	(0.036)	(0.036)	(0.069)	(0.069)	(0.066)	(0.000)	(0.020)
Prop. VP in $week_{t-1}$	-0.461^{***}	-0.458^{***}	-0.461^{***}	-0.335^{***}	-0.497^{***}	-0.367^{***}	-0.191^{***}	-0.556^{***}	-0.492^{***}
	(0.051)	(0.051)	(0.051)	(0.056)	(0.044)	(0.046)	(0.055)	(0.045)	(0.021)
Y_{t-1}							-0.211^{***}		
							(0.024)		
Mild rep. NVP in $week_{t-1}$		0.100^{**}				0.108^{**}	0.110^{***}	0.029^{*}	0.033^{*}
		(0.037)				(0.034)	(0.031)	(0.013)	(0.014)
Harsh rep. NVP in $week_{t-1}$			0.001			0.037	0.027		
			(0.053)			(0.028)	(0.025)		
Rep. VP in $week_{t-1}$				-0.364^{***}		-0.366^{***}	-0.377^{***}	-0.365^{***}	-0.385^{***}
				(0.050)		(0.045)	(0.044)	(0.034)	(0.024)
Prop. org. NVP					-0.263^{***}	-0.263^{***}	-0.247^{**}	-0.245^{**}	-0.255^{***}
					(0.074)	(0.074)	(0.079)	(0.078)	(0.016)
Observations	20,232	20,232	20,232	20,232	20,232	20,232	14,453	20,232	20,232
Adjusted R^2	0.123	0.125	0.123	0.140	0.235	0.255	0.303	0.374	0.328
Δ Adjusted R^2		0.002	0.000	0.017	0.112	0.132	0.180	0.251	0.205
Fixed effects								Grid cell	Country-year

Note: $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$
Country-level clustered SEs in parentheses

5.1 Baseline models

Table 3 shows the baseline models which includes only the covariates of interest: repression of nonviolent protest, repression of violent protest, and level of interest. Models 2-5 estimate the associations between these three variables individually on the violent escalation of protest. In model 6, I estimate the full baseline model, controlling for the proportion of violent protests. Model 7 is the baseline autoregressive model, which includes a one-week lag of the dependent variable. Model 8 is estimated using grid-cell fixed effects. These models provide a benchmark to compare the explanatory power of the variables included in the hypotheses to the extended models with controls.

Because the dependent variable measures change, coefficients with a negative sign indicate that the variable is associated with a *decrease* in the proportion of violent events, while positive coefficients indicate an *increase*. Violent escalation and de-escalation are relative concepts, and the interpretation of the coefficients are not to be conflated with an increase in the overall frequency of protest. In other words, where an increase of violent events are observed, this is *not* due to a sheer increase in the number of protest events, but rather that violent protest makes up a greater proportion of the overall protest activity.

The proportion of violent protests (relative to the total number of events) in the previous week is held constant in all models. This is because it is conceivable that repression may have a different impact in very violent environments versus nonviolent environments. Furthermore, repression (of violent protest) may be more likely in areas where there is a higher proportion of violent protest. The repression variables and the control variable are measured at the same point in time. The results therefore holds the level of violence within the grid cell constant at the time of repression, allowing us to assess the controlled association between repression in week$_{t-1}$ and violent escalation in week$_t$. The control variable is statistically significant with a negative sign, indicating that increasing levels of violence is more likely in relatively nonviolent grid cells. Grid cells where all ongoing protests are violent cannot turn *more* violent, so this result is perhaps unsurprising. Yet, the negative sign of the coefficient is theoretically interesting, as it questions the concept of escalation. I return to this in the discussion below.

With respect to the three repression variables, the results across models 1-6 in Table 3 are largely consistent. Initially, the results lend support to the two substitution hypotheses, $H_{1.1}$ and $H_{1.2}$. Mild repression of nonviolent protest in week$_{t-1}$ is positively assiociated with an increase in the level of violence in week$_t$. This result is statistically significant on the 5 % level. When the government represses nonviolent protest, using milder methods such as crowd dispersal, the results indicate that the protesters substitute nonviolence for increasing levels of violence in the next week. The association between harsh repression of nonviolent protesters is not significantly different from zero in either model. If, however, the government represses violent protest in week$_{t-1}$, the results from models 4 and 6 indicate that the protesters substitute violence for nonviolence in week$_t$, in line with $H_{1.2}$. These results hold when controlling for the share of violent events in the previous week as well as for the proportion of organized protests and the other forms of repression.

However, it is worth noting that the adjusted R^2 increases only marginally when adding each repression variable. R^2_{adj} is a measure of the amount of variance in the dependent variable that is accounted for by the independent variable(s), adjusted for the number of predictors that are added to the model (Christophersen 2013). Unlike the regular R^2, the statistic penalizes the inclusion of variables that does not significantly improve the explained variation. It should be noted that this is also the only thing that R^2_{adj} can tell us about the model specification; the statistic does not convey any causal relationships between the independent and dependent variable or whether the variables are substantially relevant (Hanck et al. 2019). The second line from the bottom in Table 3 shows the change in Adjusted R^2 from model 1. This is indicative of whether adding additional variables improve the model's explanatory power with respect to the variation of the dependent variable. As we can see, this lack of explanatory power is particularly the case for harsh repression of nonviolent protest. The coefficient is not only insignificant on the 5 % level, so that the null hypothesis that severe repression of nonviolent protest is not associated with violent escalation can therefore not be rejected; including the variable does not lead to a change R^2_{adj}. With respect to the repression variables, the greatest increase of the Adjusted R^2 stems from repression of violent protest. The two variables in model 4 account for 14 % of the observed variance in the dependent variable.

Turning to the level of organization, this is found to improve the share of the variance in the dependent variable that is modeled by 11.2 percentage points (roughly 14 percent). Substantively, the results are consistent across models 5 and 6. The coefficient has the expected negative sign, and is statistically significant on the 5 % level. The results indicate that the higher proportions of organized protests covary with decreasing levels of violence. This finding supports H_2, indicating that the higher the level of organization "on the ground", the higher the likelihood that the protesters are able to maintain nonviolent discipline. As argued above, this is likely because people participating in organized protests face the risk of negative sanctions if they turn to violence. These results hold both when controlling for repression and the level of violent protest in the previous week.

I now turn to model 7, the autoregressive model. This parsimoneous model explains approximately 30 percent of the variance in the dependent variable. In other words, controlling for the lagged dependent variable improves the model fit in comparison to the other models reported in Table 3. Adding the lagged dependent variable not only accounts for the temporal autocorrelation in the data, it also directly models the change from week_{t-1} to week_t, as it controls for omitted variables prior to week_{t-1}. The control is substantially relevant. For example, if there is a big increase in the proportion of violent events from one week to the next, it is conceivable that this leads to changes in the repression predictors. Specifically, violent escalation in week_{t-1} may lead to increased repression of violent protest in the same week, as these represent a greater threat to the government. The coefficient of the lagged dependent variable is substantially relevant if behavior at a point in time can be influenced by similar behavior at a previous point in time (Christophersen 2013). Here, we can see that violent escalation in week_{t-1} is associated with a signficant de-escalation in week_{t-1} given by the negative and statistically signficant coefficient. Similarly to the negative sign of the control for proportion of violent events in week$_{t-1}$, this too questions violent *escalation*, a point I return to below.

As we can see, the association between mild forms of repression of nonviolent protest and violent escalation remains positive and statistically significant in model 7. This provides additional support for $H_{1.1}$, stipulating that nonviolent protesters substitute nonviolence for violence in the face of repression. However, this association is only found for the mild form of repression of nonviolent protest. As in model 6, the use of harsh repressive means against nonviolent protesters is statistically insignificant.

$H_{1.2}$, where I hypothesized that repression of violent protest leads to an increase in the proportion of nonviolent protest, is also supported by the autoregressive model. When including the lagged dependent variable into the equation, the coefficient of repression of violent protest remains negative and statistically significant, as in the baseline model. In line with the hypothesis, the negative sign of the coefficient indicates that repression of violent protests is related to a shift toward nonviolence.

Turning to the level of protest organization, the results of model 7 is in line with H_2, which stated that higher levels of organization is associated with a lower risk of violent escalation. The negative sign indicates that protest-weeks where a higher proportion of the nonviolent protests are organized are at less risk of experiencing increasing violence. This association is statistically significant on the 5 % level.

Models 8 and 9 are estimated using fixed effects on the grid cell and the country-year. In model 8, a separate dummy variable for each grid cell is added, which increases the explanatory power of the model in terms of the adjusted R^2 substantially. In model 9. dummies for the year and country in which the protest occurred are added. By holding unit-specific static variables constant, these models are conservative tests of the variables of interest. An example of a country-year fixed effect is democracy. Substantively, the coefficients in models 8 and 9 are interpreted as a within-unit increase of the variable of interest. With regards to repression of violent protest and organization, the coefficients are largely similar to the other models discussed, further supporting $H_{1.2}$ and H_2. When comparing models 8 and 9 to models 6-7, slight changes to the coefficient of mild repression of nonviolent protest are observed. While still statistically significant, the reduction in the value of the coefficient indicates that the fixed effects correlate positively with both mild repression and violent escalation (Angrist and Pischke 2009, 46). I return to the strength of the associations in the data below.

Table 4: Full models

	Dependent variable:				
	Change in proportion of violent events in week$_t$				
	AR(1)		Base		AR(1)
	(1)	(2)	(3)	(4)	(5)
Constant	0.198**	0.256*	0.666**	0.693**	0.682***
	(0.066)	(0.110)	(0.216)	(0.222)	(0.198)
Mild rep. NVP$^\diamond$	0.110***		0.071***	0.068***	0.067***
	(0.031)		(0.015)	(0.015)	(0.019)
Rep. VP$^\diamond$	−0.376***		−0.373***	−0.371***	−0.383***
	(0.044)		(0.039)	(0.038)	(0.039)
Prop. org. NVP	−0.247**		−0.254***	−0.252***	−0.241**
	(0.079)		(0.074)	(0.073)	(0.078)
Prop. VP$^\diamond$	−0.191***		−0.415***	−0.412***	−0.263***
	(0.055)		(0.059)	(0.059)	(0.071)
Y_{t-1}	−0.211***				−0.179***
	(0.024)				(0.027)
Avg. travel time		0.007	−0.012	−0.015	−0.013
		(0.006)	(0.008)	(0.009)	(0.007)
Lib. dem.		0.020	−0.430	−0.402	−0.456
		(0.248)	(0.335)	(0.311)	(0.250)
Lib. dem.^{2}		0.026	0.668	0.624	0.640
		(0.363)	(0.519)	(0.483)	(0.407)
GDP/cap. (log)		−0.039**	−0.054*	−0.055*	−0.055*
		(0.013)	(0.027)	(0.027)	(0.025)
Leader tenure		0.0001	0.0001	0.00005	0.00002
		(0.0001)	(0.0001)	(0.0001)	(0.0001)
Elected		0.017	0.023	0.022	0.016
		(0.013)	(0.021)	(0.022)	(0.018)
Unemp. rate		0.006***	0.007*	0.007*	0.006*
		(0.001)	(0.003)	(0.003)	(0.003)
Duration (log)				−0.010**	−0.008*
				(0.004)	(0.003)
Ongoing VP nb.$^\diamond$				0.002	−0.00005
				(0.011)	(0.007)
Diffusion$^\diamond$				0.003	−0.004
				(0.007)	(0.006)
Observations	14,453	20,232	20,232	20,232	14,453
Adjusted R^2	0.303	0.012	0.279	0.280	0.321
Note:					*p<0.05; **p<0.01; ***p<0.001 Country-level clustered SEs in parentheses $^\diamond$ Lagged one week

5.1.1 The political environment

So far, the regression models have not included any of the control variables that convey characteristics of the protest or structural attributes of the political environment in which the protest activity occurs. As mentioned, structural theories have taught us a great deal of patterns of political violence onset. However, one of the arguments I put forward in this thesis is that structural variables are not sufficient in explaning the dynamics that lead to the violent escalation of protest, which occur in the very short term.

Table 4 presents the results from five different OLS models that incorporate the control variables operationalized in Chapter 4.7 in various ways. Again, negative coefficients indicate that the dependent variable decreases with a scale unit increase of the variable. Increasing the value of variables with positive coefficients is associated with an increase of the dependent variable amounting to the value of the coefficient.

Model 1 is similar to the autoregressive model 7 in table 3. The only difference is that I excluded harsh repression of nonviolent protests, as this was found to have little influence on explaining variance in the dependent variable; not just in terms of statistical insignificance, but we can also see that dropping the variable renders R^2_{adj} unchanged although there are slight value changes to the remaining coefficients. In this section, this is treated as the benchmark model.

Model 2 includes only the structural variables from Chapter 4.7.1. These controls are assumed to be prior to the repression and organization variables with respect to time. Controlling for prior variables allows us to assess the total effect of repression and organization - that is, the variance in the dependent variable that repression and organization can account for (Christophersen 2013). I therefore build on Model 2 in the subsequent models. The purely structural model explains a mere 1.2 percent of the variance in the dependent variable. In other words, the violent escalation of protest is poorly explained by structural factors alone.

In model 3, I incorporate the variables of interest and control for the structural variables. Urban-rural divisions, regime type and strength, and poverty are some of the most commonly cited causes of the onset of political violence. With respect to the two repression variables and level of organization, we see that the associations from model 1 hold. Higher levels of nonviolent protests that are intervened in is significantly related to an increase in the proportion of violent protest events also when controlling for the structural characteristics of the regime. This finding provides additional support for $H_{1.1}$. However, when adding the control variables, the coefficient of mild repression of nonviolent protest decreases from 0.110 to around 0.07 in in models 3-5. This indicates a positive correlation between the control variables and the independent and dependent variable (Angrist and Pischke 2009, 46). While the association is still significant, this questions its strength. Repression of violent protest is again significantly and negatively associated with the dependent variable, supporting $H_{1.2}$. With respect to controlled association between the level of organization and violent escalation, the negative and statistically signficant result provide further support for H_2.

In model 4, I have estimated the full model that controls for all variables operationalized in Chapter 4. The difference between Model 4 and the previous models is that the protest-cycle variables are added to the model. The two spatial variables, diffusion and protest in neighboring grid cells, are not statistically significant on the 5 % level. Furthermore, the inclusion of the three additional variables only led to an increase of adjusted R^2 by 0.001 in comparison to model 3. These are therefore dropped from the models presented in the remainder of this thesis. With respect to the variables of interest, the results from the previous models hold also in this model.

Model 5 is the autoregressive model with control variables. The direction of the association between the variables of interest and the dependent variable remains the same when controlling for the suggested confounders. In Figure 7, I visualize the key coefficients along with their associated confidence intervals on the 95 % level, calculated using the country-level clustered standard errors. If the confidence interval crosses the dashed zero-line, the effect of the variable is not statistically significant on the 5 % level. The figure compares the base model and the autoregressive model, with and without control variables. As we can see from the figure and when comparing models 1 and 5 in Table 4, controlling for attributes of the political environment and protest duration does not change the substantive interpretation of the variables of interest, in terms of direction and significance of the relationship. In addition, the values of the coefficients remain largely the same, indicating only a slight correlation between the control variables and the dependent and independent variables of interest. As I have shown, higher levels of (mild) repression of nonviolent protest is associated with an increase in the proportion of protest violence in the following week. The association between repression of violent protest and violent escalation is also robust across the four models. The patterns in the data show that repression of violent protest is associated with decreasing levels of violence in the following week. Providing support to H_2, the models have also unveiled an association between higher levels of organized protest and decreasing levels of violence.

Table 4 illustrates some key points of the structural variables. First and foremost, in all models 2-5, the structural variables are generally insignificant on the 5 % level when clustering the standard errors on the country-level. The only exception is GDP per capita, which has the expected negative sign, and unemployment rate, where the coefficient is positive, as expected, but low (Gustafson 2019). In other words, poverty and high unemployment rates are both associated with violent escalation of protest. To exemplify one reason for this lack of explanatory power across the structural control variables, a country's (i.e. a specific grid cell's) democracy score does not change from one week to the next, except from between week 52 and week 1 in a new year, because it is measured in yearly increments. In addition, large changes from one year to the next is rare. With respect to whether structural attributes of the political system are of importance when assessing risk of violent escalation, no clear conclusions can be drawn from this discussion. Because of the lack of sufficiently disaggregated data, both temporally and spatially, there is little variation within units over time. Using these to estimate changes in the dependent variable is largely futile when the dependent variable is subject to short-term variation.

The lagged dependent variable controls for all omitted variables prior to week$_{t-1}$, but not for changes that occur from this week to week$_t$. In the autoregressive model (model 1), such structural factors are thus controlled for as omitted variables, precisely because they do not change between week$_{t-1}$ and week$_t$. This illustrates the fundamental difference between temporally aggregated and disaggregated models, as well as why I choose to rely on the autoregressive model when interpreting the results[11]. Also note that I do not report a model with the structural control variables using Fixed Effects. The grid-cell dummy variables in model 8 in Table 4 similarly controls for unit-specific factors, which renders the coefficients of the control variables statistically insignificant (demonstrated in Appendix). In section 5.5, I return to the control variables in a discussion of the causal interpretation of the models.

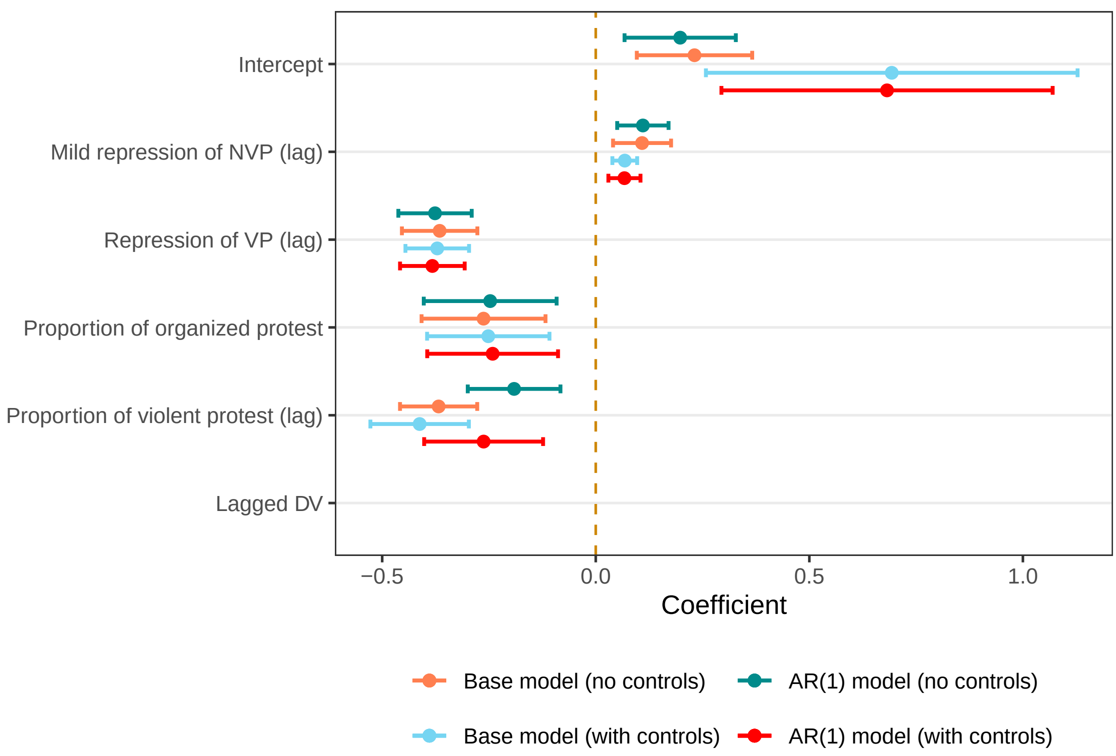

Figure 8: Coefficient estimates

[11]I recognize that estimating a multilevel model could provide further insight into the relationship between level-1 and level-2 variables and the dependent variable. However, when attempting this in R, the model failed to converge when adding level-2 variables. A hierarchical model with a random intercept by country, without level-2 variables produced the same results with respect to repression and organization as the models described here.

5.2 Interaction: Organization and repression

H_3 hypothesizes that the level of organization moderates the effect of repression on violent escalation. In order to test this proposition, I add two interaction terms to the linear model, thereby assessing the conditional relationship between organization, repression and violent escalation of protest. Here too, I present the results from various specifications of the model. Table 5 shows the baseline model estimated without corrections, the first-order autoregressive model and the models with fixed effects on grid cell and country year. The control variables are included in model 2 for comparative reasons, but based on the above discussion I will not assess these in depth. As for the models in Table 4, harsh repression of nonviolent protest is excluded from this model (included in Appendix).

Across all five models, the coefficients of both interaction terms are negative and statistically significant on the 5 % level, which implies that the interaction terms should be included. The OLS model is written as:

$$\tilde{Y} = \beta_0 + \hat{\beta}_1 repression_{NVP} + \hat{\beta}_2 repression_{VP} + \hat{\beta}_3 organization + \hat{\beta}_4 rep_{NVP} \times org + \hat{\beta}_5 rep_{VP} \times org + \hat{\beta}_k X_k + e_i,$$

where $\beta_k X_k$ symbolize the control variables and their estimated coefficients, including the lagged Y_{t-1} or the grid-cell fixed effects.

Because both organization and repression are continuous variables, the interpretation of the coefficients is somewhat more complex than in instances where one of the constituent variables represents the absence or presence of a condition. However, I begin with this parsimoneous interpretation to exemplify. In an instance where no protests are organized, i.e. when the organization variable equals zero, the equation above can be rewritten as:

$$\tilde{Y} = \beta_0 + \hat{\beta}_1 repression_{NVP} + \hat{\beta}_2 repression_{VP} + \hat{\beta}_3 0 + \hat{\beta}_4 rep_{NVP} \times 0 + \hat{\beta}_5 rep_{VP} \times 0 + \hat{\beta}_k X_k + e_i,$$

Table 5: Interaction models

	Dependent variable:				
	Change in proportion of violent events in week$_t$				
	Base		AR(1)	FE	
	(1)	(2)	(3)	(4)	(5)
Constant	0.221** (0.068)	0.683** (0.221)	0.185** (0.064)	−0.000 (0.000)	0.194*** (0.035)
Mild rep. NVP$^\diamond$	0.194** (0.070)	0.149** (0.047)	0.219** (0.078)	0.123** (0.039)	0.118** (0.043)
Rep. VP$^\diamond$	−0.197** (0.076)	−0.200** (0.066)	−0.212** (0.074)	−0.218*** (0.061)	−0.218*** (0.062)
Prop. org. NVP	−0.241*** (0.071)	−0.230*** (0.069)	−0.221** (0.075)	−0.225** (0.074)	−0.234** (0.074)
Prop. VP$^\diamond$	−0.366*** (0.047)	−0.410*** (0.059)	−0.190*** (0.055)	−0.553*** (0.045)	−0.490*** (0.040)
Avg. travel time		−0.016 (0.009)			
Lib. dem.		−0.411 (0.314)			
Lib. dem.^{2}		0.637 (0.486)			
GDP/cap. (log)		−0.054* (0.027)			
Leader tenure		0.00004 (0.0001)			
Elected		0.020 (0.022)			
Unemp. rate		0.007* (0.003)			
Duration (log)		−0.010** (0.003)			
Ongoing VP nb.$^\diamond$		0.001 (0.011)			
Diffusion$^\diamond$		0.003 (0.007)			
Y_{t-1}			−0.210*** (0.022)		
Mild rep. (NVP) × Prop. org.	−0.175* (0.072)	−0.167* (0.071)	−0.204* (0.088)	−0.191** (0.066)	−0.175* (0.072)
Rep. (VP) × Prop. org.	−0.445*** (0.098)	−0.450*** (0.097)	−0.429*** (0.090)	−0.381*** (0.089)	−0.440*** (0.090)
Observations	20,232	20,232	14,453	20,232	20,232
Adjusted R^2	0.264	0.289	0.313	0.381	0.336
Fixed effects				Grid cell	Country-year

Note: *p<0.05; **p<0.01; ***p<0.001
Country-level clustered SEs in parentheses

In such cases, the estimated association between an increase in mild repression of nonviolent protest and violent escalation amounts to $\hat{\beta}_1$, or 0.219 in the autoregressive model. Similarly, the estimated association between a unit increase in repression of violent protest and violent escalation is $\hat{\beta}_2$, or a decrease in the proportion of violent events of -0.212 from week$_{t-1}$ to week$_t$ when relying on the AR(1) model. The corresponding coefficients in model 4, estimated with grid-cell fixed effects, are 0.123 for mild repression of nonviolent protest and -0.218 for repression of violent protest. This specific situation where organization equals zero is the case for roughly 38 percent of the observations in the full dataset, and 67 percent of the observations in the data used to estimate the autoregressive model[12]. Similarly, where no repression of neither nonviolent nor violent protest occurred, the association between level of organization and violent escalation is simply $\hat{\beta}_3$. A unit increase in organization is associated with a decrease of the proportion of violent events amounting to -0.221 (model 3) and -0.225 (model 4).

As we can see, the results presented in Tables 3-4, with respect to the three variables of interest, hold also when introduction the interaction terms across all four models in Table 5. Turning first to $H_{1.1}$, mild repression of nonviolent protest remains a substantively strong and statistically significant predictor of changes toward more violence, when organization equals zero. The same conclusion is reached with regard to repression of violent protest, which also remains negatively associated with violent escalation of protest, providing additional support for $H_{1.2}$. The conditional association between level of organization and the dependent variable is also negative and statistically significant across the three interaction models, in line with H_2, when the two repression variables equal zero.

As the interaction variables are statistically significant, it is safe to conclude the association between repression of either kind and violent escalation depends on the level of organization. Conversely, the association between the proportion of organized protests and violent escalation depends on the level of repression in the preceding week. In simple terms, with increasing repression of either nonviolent or violent protest, the proportion of violent events in week$_t$ decreases, as the level of or-

[12]That is, where the first observation is excluded in order to lag the dependent variable.

ganization increases. In Figures 9-12 I visualize the slope of the repression variables across different values of the organization variable, which ranges from 0 to 1. These plots are an effective way of inspecting the interaction terms (Brambor, Clark, and Golder 2006). The plots are based on the autoregressive and the fixed effects models. In the plots, the shaded area represents the 95 % confidence intervals, calculated using the country-clustered standard errors. If the confidence band crosses zero on a value of organization, the association between repression and violent escalation is statistically indistinguishable from zero. The plots are drawn using the correction suggested by Esarey and Sumner (2017), which controls for false positive rates, providing a more conservative test of the conditional relationships in the data.

Figure 9 shows that the marginal association between mild repression of nonviolent protest and violent escalation decreases the more protests are organized, based on the estimates in model 3, Table 5. Here, we see that the slope, i.e. the coefficient, of mild repression of nonviolent protest is contingent upon the value of organization variable. This means that when the government represses nonviolent protest in places where a low proportion of protests are organized, the association between this type of repression and violent escalation is stronger than where there is a high level of organized protests. Nonetheless, the confindence interval is large on low levels of organization, reflecting the degree of uncertainty of the coefficient estimate. Furthermore, at increasing levels of organization, the slope draws closer to zero, i.e. no effect. In fact, at very high levels of organization, the effect is statistically indistinguishable from zero at the 5 % level, as the 95 % confidence intervals include zero.

In figure 10, the slope of repression of violent protest is plotted against different level of organization. Even at low levels of organization, repression in week$_{t-1}$ is associated with a decrease of violent activities in week$_t$. At higher levels of organization, this association grows stronger. In other words, where there is a high level of organization, repression of violent protest is likely to lead to a decrease in violent events in the preceding week. The repression coefficient is statistically significant across all levels of organization.

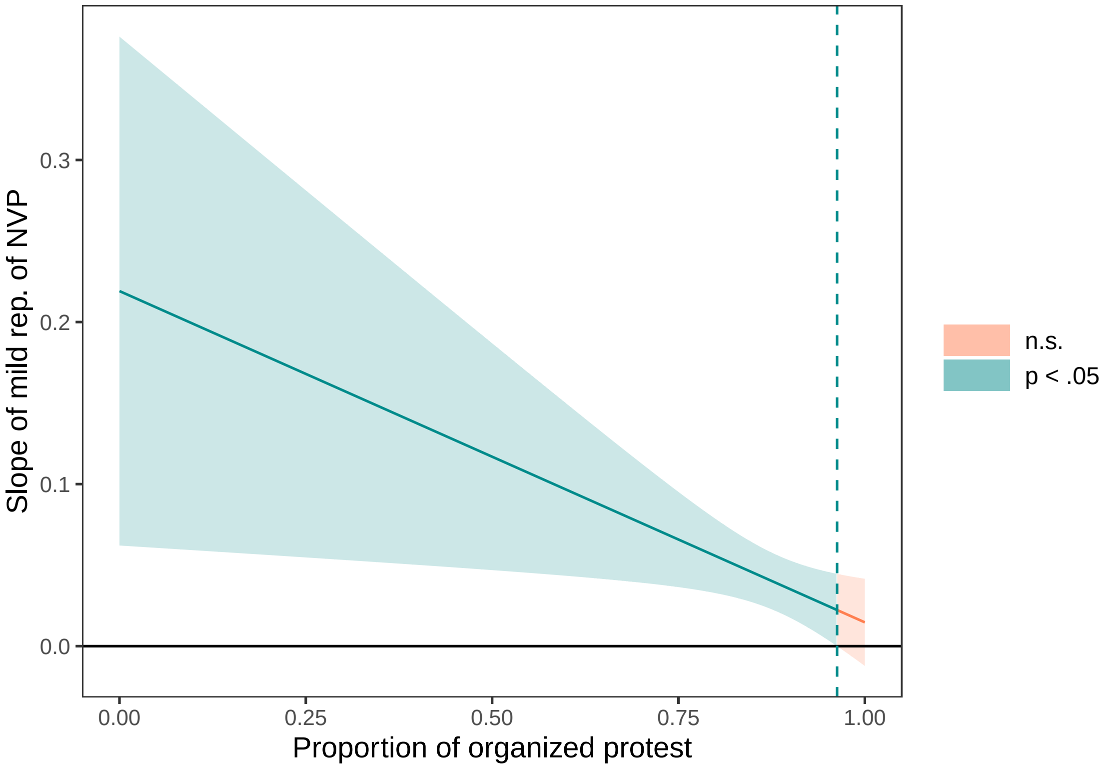

Figure 9: Interaction plot of mild repression of NVP in week$_{t-1}$ and proportion of organized protests, AR(1)

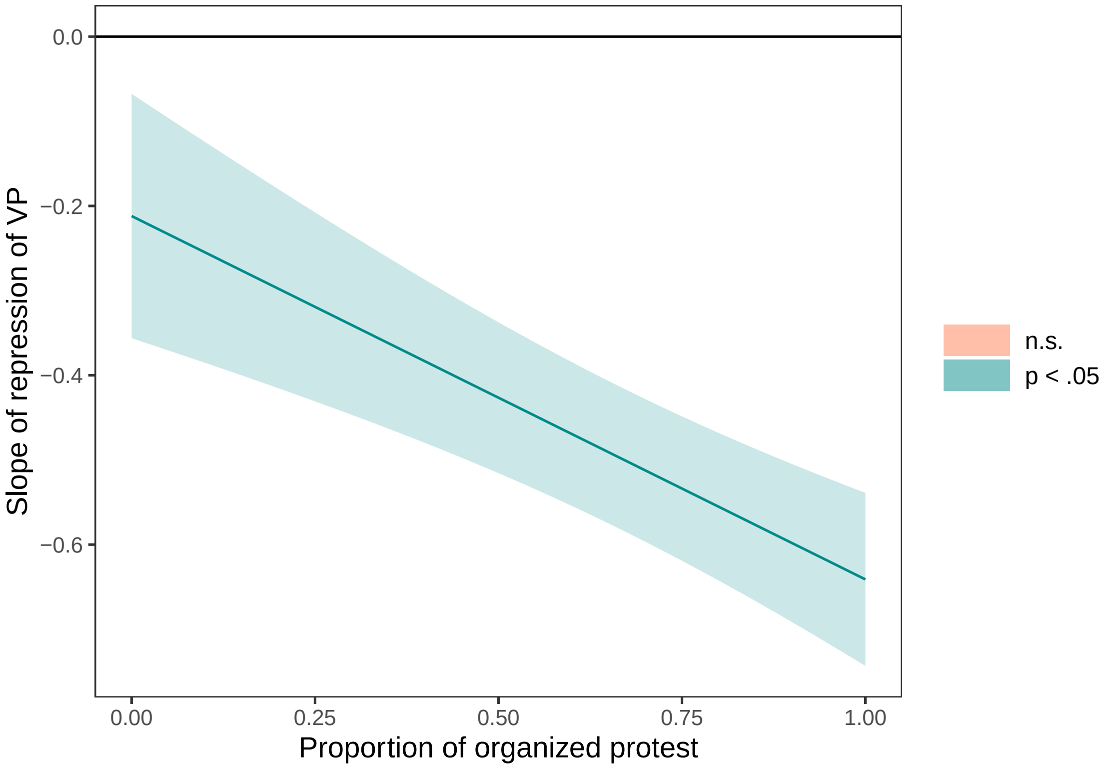

Figure 10: Interaction plot of repression of VP in week$_{t-1}$ and proportion of organized protests, AR(1)

Figures 11 and 12 are based on model 4, Table 5, which is estimated using grid-cell fixed effects. In broad terms, the associations in the data hold when comparing to the autoregressive model. Turning first to figure 11, one important difference needs to be pointed out. As we can see, the slope of mild repression of nonviolent protest decreases with increasing levels of organization. However, where around 50 % or more of the protests are organized, the conditional association between mild repression of nonviolent protest and violent escalation is statistically insignificant. This is a much lower threshold than in the autoregressive model. The results are therefore mixed when it comes to higher levels of organization. What these plots do convey, however, is a relatively robust association between low to medium levels of organization, repression of nonviolent protests and violent escalation of protest. When faced with repression, the lack of organization seemingly leads to lower nonviolent discipline in nonviolent protests.

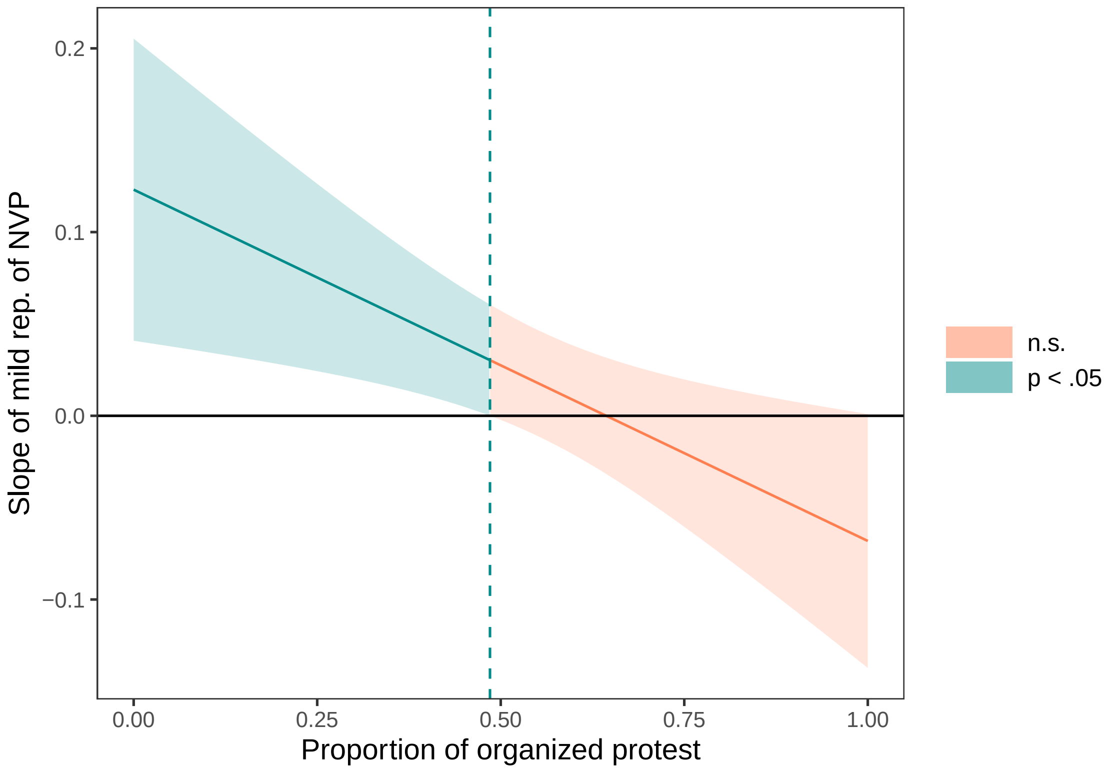

Figure 11: Interaction plot of mild repression of NVP in week$_{t-1}$ and proportion of organized protests, FE

Figure 12, based on the fixed effect model, is largely similar to figure 10, despite the different modeling strategies. Here too, high levels of organization is associated with a substantial decrease of the association between repression of violent protest and violent escalation. The repression slope is significant across all levels of organization. To exemplify, in areas where all protests are organized, increasing repression of violent protest is associated with a reduction of violent events in week$_t$ compared to the previous week. This reduction is greater than where no protests are organized, although the proportion of violent events decreases also here.

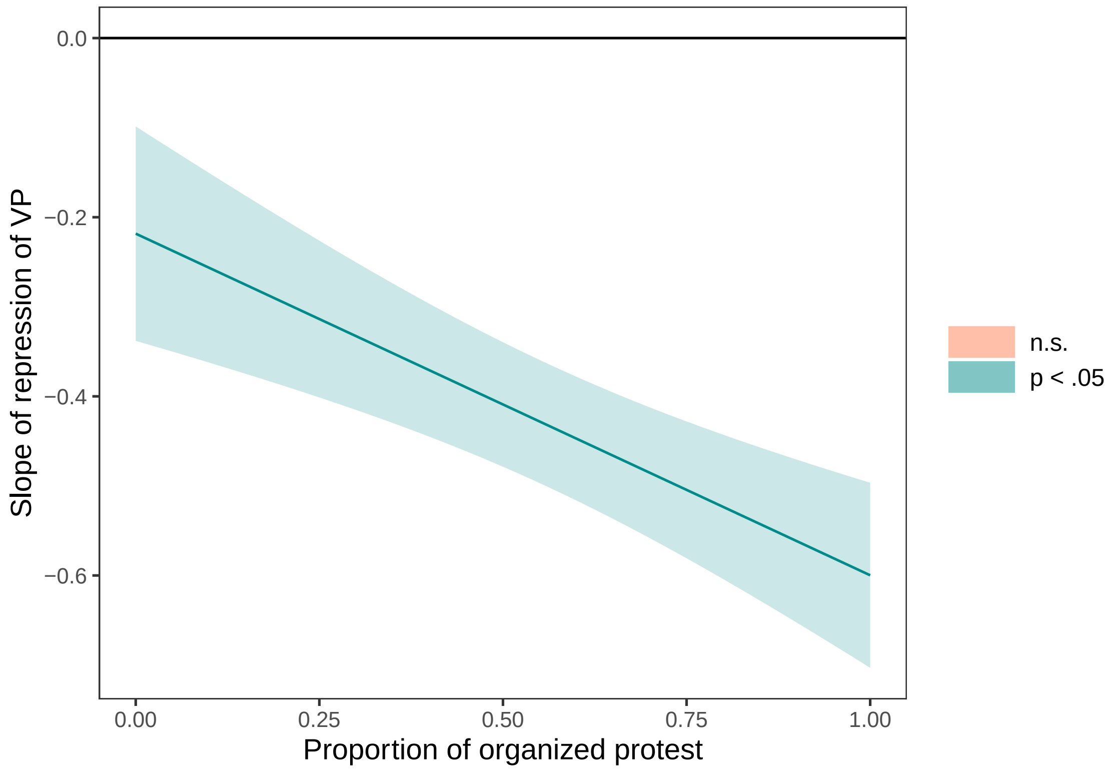

Figure 12: Interaction plot of repression of VP in week$_{t-1}$ and proportion of organized protests, FE

The interaction effect can of course also be viewed from the opposite perspective, by examining the slope of organization across different levels of repression. The negative coefficient of the interaction term indicates that also the association between organization and violent escalation decreases with increasing levels of repression. The substantive interpretation is thus somewhat different from stating that organization reduces the escalatory impact of repression. First, where there is no repression, i.e. when both repression variables equal zero, the association between organization and violent escalation is negative across all model specifications. This follows from the equation above. In other words, in the absence of repression, increasing the organization variable is associated with decreases in violent events - or as a de-escalatory factor. Again, the organization coefficient is plotted against different levels of repression of nonviolent and violent protest, respectively, in figures 13 and 14. As there were no substantive differences between the autoregressive and the fixed effects model, I only report the results from the AR(1) model (fixed effects included in Appendix).

As we can see from figure 13, the slope of organization decreases when more nonviolent protests are met with mild repression. The coefficient is significant on the 95 % level across all levels of mild repression. In substantive terms, this indicates that the de-escalatory impact of organization is stronger when nonviolent protesters face mild repression, compared to when there is no use of repressive measures against nonviolent protesters. The same conclusion is reached when inspecting figure 14. While organization is associated with a decrease in the proportion of violent protests when violent protests are not facing repression, the association grows stronger when a substantial proportion of the violent protests in week$_{t-1}$ met repressive means. To summarize, in order to avoid violent escalation, organization seems to be of importance, particularly in the face of repression.

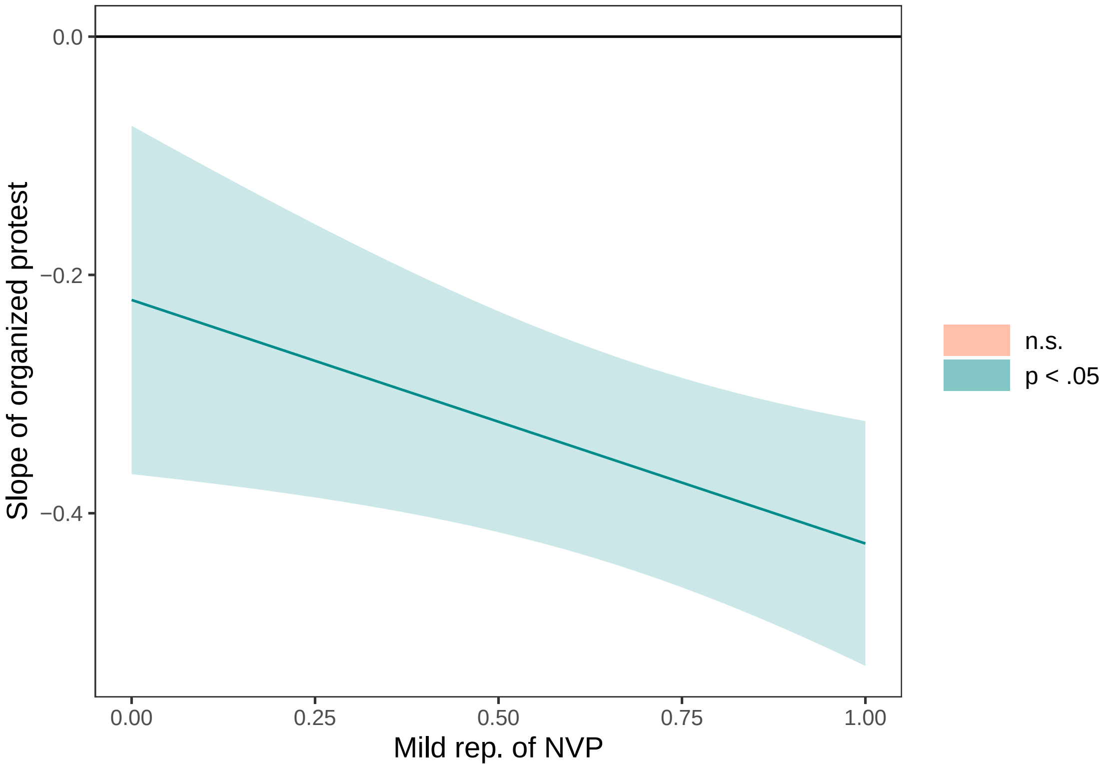

Figure 13: Interaction plot of proportion of organized protests and mild repression of NVP in week$_{t-1}$, AR(1)

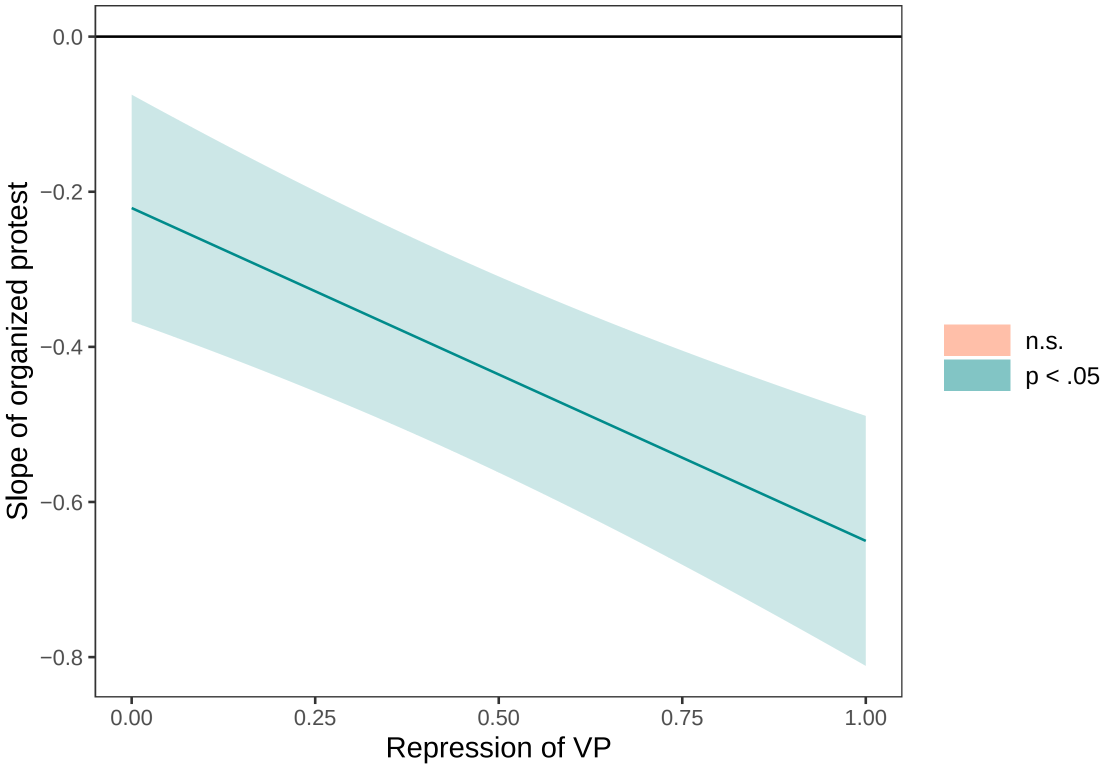

Figure 14: Interaction plot of proportion of organized protests and repression of VP in week$_{t-1}$, AR(1)

5.3 The strenght of the relationships

Up until this point, most focus has been placed on assessing the signs of the regression coefficients. The strength of the associations, in terms of the absolute values of the coefficients, also need to be addressed. Substantively, coefficients are interpreted as the change in the dependent variable corresponding with a unit increase in the independent variable, controlled for the other variables included in the regression. The coefficients need to be interpreted in terms of the scale of the dependent variable, which ranges from -1 to 1, and the scale of the three variables of interest, ranging from 0 to 1. These proportional scales make the interpretation of a unit increase somewhat complicated. Therefore, the analysis above is largely focused upon the direction rather than the magnitude of the association.

One way to assess the strength of the relationships is to use the standard deviation of the dependent variable. From Table 2, we see that the dependent variable in the full sample has a mean of 0.058 and a standard deviation of 0.339. In other words, in the full sample, the average deviation from the mean is 0.339. However, the autoregressive models are estimated using a smaller sample. Table 6 summarizes the mean and standard deviation of the dependent variable in the two samples used in the above analysis.

Table 6: Dependent variable

Statistic	N	Min	Max	Mean	St. Dev.
Full sample	20,232	−1	1	0.058	0.339
AR(1) sample	14,453	−1.000	1.000	0.029	0.332

In the following, I use the results from the autoregressive and fixed effects models reported in Table 5 as the interaction terms were statistically significant and therefore should be modeled. I only consider the coefficients of the two repression variables and the organization variable. The coefficients of the interaction terms are not meaningful to interpret on their own. The values in Table 7 are calculated by dividing the respective coefficients (Table 5) on the standard deviation of the sample used to estimate the regression, multiplied by 100. These interpretations hold only when the other variables of interest equal zero, due to the interaction terms.

I first turn to the mild repression of nonviolent protest. In the autoregressive model, one unit increase of the the variable corresponds to a 0.219 increase of the proportion of violent events, when no protests are organized. This amounts to 66.1 percent of the dependent variable's standard deviation. In the grid-cell fixed effects model, the coefficient decreases to 0.123, and a unit increase of the independent variable in amounts to 36.3 percent of the standard deviation of the dependent variable. Repression of violent protest yields similar results. Here, the coefficient is negative in both models. In the autoregressive model, the respective coefficient is -0.212. As the standard deviation is the deviation from the mean, I divided this by the negative standard deviation. Thus, an increase of repression of violent protest amounts to a decrease in the dependent variable of -0.212, which constitutes 63.9 percent of the standard deviation. For the fixed effects model, the corresponding value is 64.4 percent. Finally, the coefficient of organization is -0.221 and -0.225 in the autoregressive and fixed effects models, respectively. This amounts to 66.7 percent of the dependent variable's standard deviation in the autoregressive model, and 66.5 percent in the fixed effects model.

In sum, for all three variables, the coefficient is roughly 65 percent of the standard deviation of Y, apart from mild repression of violent protest in the fixed effects model. Because the standard deviation is a measure of the average deviation from the mean in the sample, the relationships in the data can be said to be relatively strong.

Table 7: Strength of associations

Variable	Strength	Direction
Mild repression NVP, AR(1)	66.1	Positive
Mild repression NVP, FE	36.3	Positive
Repression VP, AR(1)	63.9	Negative
Repression VP, FE	64.4	Negative
Organized NVP, AR(1)	66.7	Negative
Organized NVP, FE	66.5	Negative

5.4 Validity of the results

I now turn to a discussion of whether statistical inferences can be drawn from the above presented models.

Whether statistical conclusions are valid depends in part on whether the associations are strong, which has been discussed in the previous section, and statistically significant at conventional levels of uncertainty, typically the 5 % level which is what I have adopted in this thesis (Lund 2002). There is therefore a 5 % risk that one or more of the results above are only products of chance and that Type I-errors have been made. Nonetheless, this is the conventional level of uncertainty applied in the literature. With regards to this point, I have shown that the relationships in the data are largely consistent in terms of the significance and form of the controlled associations between repression, organization and violent escalation of protest, using various specifications of the linear model. These statistical results lend support to the substantive hypotheses.

A related aspect of internal validity is that the coefficients are unbiased and consistent estimates of the true coefficient (Stock and Watson 2007, 313). Although the true effect of a unit increase in repression or organization on violent escalation is unknown, this has been assessed with a number of robustness checks. I now turn to discuss the results from these tests. The regression tables are included in Appendix.

A number of influential observations that have extreme values on the independent variables, and thereby a large residual, were detected. As non-repression in week$_{t-1}$ is more common than the occurrence of repression, given by the low mean of the repression variables (see Table 2), those observations that have high values on these variables (closer to 1) may be likely to drive the results; which is not particularly surprising. The same goes for the dependent variable. The analysis above not only models the instances of violent escalation of protest, or the consequences of repression, but also the instances of non-escalation where the dependent variable equals zero and the instances of non-repression. When modeling only escalation, where only observations with a value on the dependent variable greater than zero - selection on the dependent variable is problematic in itself - the substantive results from the analysis above prove to be generally robust (see Appendix). In order to test whether the coefficients are biased because of observations with high leverage or outliers, I estimated a robust regression model. This showed that the coefficients, in terms of direction, are robust to this method of estimation. However, the coefficient estimates are much lower for most of the variables - except repression of violent protest, where the coefficient is several times higher. In sum, the coefficient estimates reported above are robust in terms of the direction of the association, but the strength of the association may be biased.

As a robustness check of whether the operationalization of the independent and dependent variables as proportions drove the results, I estimated models with various specifications of binary independent and dependent variables. The OLS models using dichotomous predictor variables[13] also show that the substantive results are generally robust when modeling only if repression of either kind occurred or if organization of nonviolent protest was present. However, when assessing the interaction term, the results are somewhat inconsistent. In the AR(1) model, the association between repression of violent protest and violent escalation is statistically insignificant where no protests are organized, but negative and statistically significant where protest organization is present. In the fixed effects model, the association between mild repression of nonviolent protest and the dependent variable is insignificant where one or more protests are organized, but positive and significant where no organization is present. Taken together, these results highlight the de-escalatory impact of organization.

A second threat to statistical conclusion validity is more prevalent in the above analysis. Although I have used various specifications of the linear model, there is a risk that the observed relationships in the data is due to the choice of the linear model and the acceptance of the OLS assumptions. With a metric dependent variable, OLS regression is the Best Linear Unbiased Estimator (BLUE), given that the data fulfills the assumptions underlying the linear model (Christophersen 2013). The assumptions that underlie the OLS model particularly concern the residual term, $e_i = Y_i - \tilde{Y}$.

[13]Coded 1 if the original variable is greater than 0, and 0 otherwise.

The residual is assumed to be normally distributed, homoscedastic and without autocorrelation (Christophersen 2013). In the OLS models presented above, the residuals are neither normally distributed nor homoscedastic. According to Christophersen (2013), non-normally distributed error terms are of most concern in small sample sizes. Although the data is clustered, the effective number of observations is still rather high. Heteroscedastic residuals typically indicate skewed variables, outliers, or units with high leverage, the latter addressed above. According to Worrall (2010), problems of heteroscedasticity typically increase when data is collected across time and space, which is the case for my data. These issues of heteroscedastic and non-normal residuals do not affect the coefficient estimates when running OLS regression. However, it can produce biased standard errors. As elaborated upon above, precise standard errors are crucial in order to draw inferences based on the results. The results presented in the above are estimated using a range of correction measures to account for underestimation of standard errors, which greatly enhance the risk of Type I error, including clustering standard errors on the country level, introducing the lagged dependent variable, and estimating fixed effects. Overall, the clustered standard errors were several times larger than heteroskedasticity-robust standard errors (reported in Appendix), effectively reducing the risk of Type I-error. Nonetheless, the threat to statistical validity is evident in the data, and some caution should be exercised in the interpretation of the significance of the results.

While these concerns cannot be perfectly addressed, I include a logistic regression model in Appendix as an additional sensitivity tests. This is useful for assessing whether the results reported above are dependent on the choice of modeling strategy. The logistic regression show that the results are robust when using a binary dependent variable. This alleviates the concern that the results are driven by the OLS model and not by the patterns in the data. In sum, the various sensitivity tests increase the confidence in that the results reported above represent the patterns in the data, at least in terms of the direction of the relationship and statistical significance.

5.4.1 Temporal lags

Table 8: Additional temporal lags

	Dependent variable:					
	Change in proportion of violent events in week$_t$					
	OLS		AR(1)		FE	
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	0.171***	0.161**	0.148***	0.140**	0.775***	0.790***
	(0.064)	(0.063)	(0.057)	(0.056)	(0.028)	(0.043)
Interv. NVP (t-1)	0.113***	0.082***	0.093***	0.085***	0.029	0.024
	(0.027)	(0.021)	(0.024)	(0.024)	(0.019)	(0.018)
Interv. NVP (t-2)	−0.005	−0.020	0.005	0.007	−0.057**	−0.058*
	(0.035)	(0.037)	(0.036)	(0.033)	(0.025)	(0.030)
Interv. NVP (t-3)		0.062**		0.064**		−0.0001
		(0.028)		(0.032)		(0.030)
Repr. VP (t-1)	−0.386***	−0.391***	−0.357***	−0.365***	−0.355***	−0.327***
	(0.043)	(0.044)	(0.043)	(0.048)	(0.040)	(0.037)
Repr. VP (t-2)	0.219***	0.183***	0.014	−0.001	0.092***	0.072***
	(0.038)	(0.031)	(0.031)	(0.035)	(0.019)	(0.015)
Repr. VP (t-3)		0.218***		0.208***		0.113**
		(0.058)		(0.053)		(0.049)
Prop. Org. (t)	−0.256***	−0.250***	−0.252***	−0.247***	−0.235***	−0.229***
	(0.086)	(0.086)	(0.086)	(0.085)	(0.081)	(0.080)
Prop. Org. (t-1)	0.060**	0.055**	0.051***	0.050**	0.031**	0.025**
	(0.026)	(0.023)	(0.019)	(0.021)	(0.014)	(0.012)
Prop. Org. (t-2)	−0.003	−0.004	0.023***	0.025***	0.008	0.010
	(0.014)	(0.013)	(0.008)	(0.007)	(0.010)	(0.010)
Prop. Org. (t-3)		0.003		−0.003		0.010
		(0.016)		(0.015)		(0.009)
Prop. VP (t-1)	−0.350***	−0.378***	−0.170**	−0.217***	−0.590***	−0.625***
	(0.056)	(0.060)	(0.077)	(0.073)	(0.047)	(0.046)
Y_{t-2}			−0.230***			
			(0.029)			
Y_{t-3}				−0.220***		
				(0.023)		
Observations	11,626	9,886	9,886	8,694	11,626	9,886
Adjusted R^2	0.305	0.316	0.320	0.342	0.408	0.420
Note:						*p<0.1; **p<0.05; ***p<0.01

I now turn to a brief discussion of the temporally lagged independent variables and the implications of this for drawing inferences based on the results above.

The repression variables are lagged one week. Thus, the results only convey the association between repression in week$_{t-1}$ and the change in violent events in week$_t$. In other words, one cannot draw inferences on the longer-term influence of repression based on the above analysis. One of the aims of this thesis is to explore the short-term association between repression and violent escalation of protest. The immediate effect of repression, measured on a week-by-week level, has been left largely unexplored in the literature. This is partly because of the lack of sufficiently disaggregated data and partly due to the operationalization of repression as the frequency of repressive events in the recent past. As such, the results presented above are novel in their own right.

Introducing additional temporal lags effectively reduces the number of observations. In the base model, nearly half of the observations are lost when lagging the independent variables by only one additional unit. In order to test whether the association between different temporal lags of the repression and organization variables and the dependent variable, the base OLS, autoregressive, and fixed effects models with independent variables measured in week$_{t-2}$ and week$_{t-3}$ were estimated and reported in Table 8.

The results from these models indicate that the results presented above are sensitive to the chosen number of temporal lags. Mild repression of nonviolent protest in week$_{t-2}$ is not statistically significant in the OLS and AR(1) models. In the fixed effects model, this variable is associatied with a decrease in the proportion of violent events in week$_t$. Mild repression in week$_{t-3}$ show similar mixed results. A similar change-of-sign is also found for repression of violent protest. While repression in week$_{t-1]}$ is consistently associated with a decrease of violent protests in week$_t$, repression of violent protest in week$_{t-2}$ is associated with an increase in week$_t$. The same result is found when looking at repression in week$_{t-3}$, also associated with an increase of the dependent variable across all three models. In other words, the substitution hypothesis, $H_{1.1}$ and $H_{1.2}$ are supported only in the immediate term.

Turning to the organization variable, introducing temporal lags give similar mixed results. In the models above, the variable is not lagged temporally. Rather, it is thought that the level of organized protests in week$_t$ has a direct impact on whether the proportion of violent protest changes from week$_{t-1}$ to week$_t$. The more nonviolent protests that are organized in week$_t$, the higher the likelihood that nonviolent discipline can be enforced in this week. The proportion of organizations in the previous week may also excert an influence on the *change*, through its impact on the protest level in week$_t$. This, however, is captured by the level of violent protest in week$_{t-1}$. Higher levels of organization in week$_{t-1}$ is weakly associated with an increase in the proportion of violent events in week$_t$, while the additional temporal lags give mixed and largely statistically insignificant results.

From this, it is evident that the scope of the inferences to be drawn is narrow. While it must be reiterated that the additional temporal lags reduce the number of observations and clusters in the data, the only results that are consistent with the hypotheses of this thesis are (1) the level of repression in week$_{t-1}$ and (2) the level of organization in week$_t$. From a theoretical perspective, it is more plausible that temporally proximate events of repression influence violent escalation, rather than repressive events weeks later. It is also more plausible that the level of organization on the ground in a given week constrains the protesters from turning to violence than whether protests were organized in previous weeks.

5.4.2 External validity

External validity refers to the potential for generalizing inferences across units, times and situations; or from sample to population (Lund 2002, 110). The sample used in this thesis is 20232 protest-weeks in 54 countries, over the course of 9 years. With respect to existing studies on violent escalation of dissent, this is a considerably higher number of observations, both with respect to the sheer number of observations as well as the different contexts that are modeled, in terms of political environments as well as over many years.

Where Moore (1998) tested the substitution hypothesis on the activities of only two dissident organizations in two countries, I have provided additional support for the hypothesis using a much larger sample. In addition, I have explicitly modeled organization level as a variable rather than an assumption, thereby including a wider variety protests including those that do not fall into the conventional definition of a campaign.

Furthermore, the sample includes protests both in Africa and Asia. The generalizability of the results to other regions of the world depends on whether protests elsewhere are inherently different from those included in the sample and on whether the political, judicial or institutional environments in which other protests occur differ fundamentally from those studied here (Stock and Watson 2007, 314–15). The possibility of this cannot be excluded, pointing to the need for future research on protests in other parts of the world. This point is further emphasized with regards to the influence of repression, which Gustafson (2019) found not to increase the probability of violent escalation in his sample that consisted of protest events in Africa and Latin America from 1991 to 2017, although the author and I operationalized the concept differently.

5.5 Evaluating the results: Causality

An additional central assumption underlying any type of regression is that there are no omitted variables that explain the dependent variable or is correlated with the independent variable(s) (Hanck et al. 2019). Any omitted variables are included in the residual. While there is no way to perfectly control for omitted variable bias, the autoregressive models control for all omitted variables up until week$_{t-1}$ and the fixed effects models control for all omitted variables that are assumed to be static within groups.

Control variables are particularly relevant for the question of causal inferences, as it allows for controlled comparisons between different units (Angrist and Pischke 2009). The ACLED data used for estimating the regression models presented above is observational. When using observational data, rather than conducting experiments, the causal interpretation of the regression estimates should be closely evaluated, which is what I do in this section.

Causality implies that X leads to Y through a causal mechanism. With this notation, X is the cause, or an event or condition, "that *raise the probability* of some outcome occurring" (Gerring 2005, 169, emphasis in original). If the key predictor X is randomly assigned to the sample, as in experimental designs, the results from the regression model can be interpreted in causal terms (Angrist and Pischke 2009). Hence, the causal effect is defined as "the effect on an outcome of a given action or treatment, *as measured in an ideal randomized controlled experiment*" (Stock and Watson 2007, 9, italics added). However, "assigning" repression to a randomized sample of protest events in order to see whether there is a causal relationship between the predictor and violent escalation is not only unfeasible, but also unethical. Similarly, it is impossible to realistically study the effect of organization on violent escalation of protest using the experimental ideal. In the absence of an experimental design, the threat to the study's internal validity is evident.

Identifying the causal mechanism is significantly more difficult than unveiling correlations in the data, and must be done on the basis of theory. Yet, if a causal mechanism exists, a covariational association between X and Y must, inevitably, also exist (Gerring 2005). In other words, causality implies a correlation, but correlation does not necessarily imply causality. While the associations in the data are robust across model specifications, this does not necessarily mean that higher levels of repression of violent protest is a cause of violent de-escalation or that mild repression of nonviolent protest leads to violent escalation. Addressing the chicken-or-egg problem with regards to repression versus violent protest is beyond the scope of this thesis. As governments are more prone to repress violent protest, nonviolent protests may in some instances escalate in a violent direction before police forces are dispatched to control the situation. This is a level of nuance that is lost when relying on quantitative observational data, where protests are either coded as violent or nonviolent and repression either was observed or not observed. While I argue that this thesis is novel in its temporal disaggregation, events *are* aggregated to the weekly level, which inevitably implies a lack of nuance. However, for the repression variables, this potential problem of reverse causality is modeled directly through lagging the variables one week. Here, the theorized causal mechanism is that repression increases the costs of protest, leading protesters to substitute their mode of protest in the following week.

With regards to the organization variable, the endogeneity problem is not addressed in the same way as this is measured during the same week as the dependent variable. An effect cannot be prior to its cause (Gerring 2005, 175). It is therefore plausible that a decrease in the proportion of violent events leads to more nonviolent protests to mobilize, and a number of these are organized - in effect, opening a window of opportunity. However, I argue that an increase in the number of organized nonviolent protests effectively prevent these protests from turning violent, leading to the observed decrease in the proportion of violent protests relative to the previous week. This theoretical argument is further substantiated by the results when incorporating additional temporal lags. Higher levels of protest organization in week$_{t-1}$ is weakly associated with an increase in the proportion of violent events in week$_t$. However, it seems illogical to claim that the more protests are organized in the prior week is a cause of violent escalation, without taking into account that the level of organization in week$_{t-1}$ influenced the maintenance of nonviolent discipline in week$_{t-1}$. One avenue for futher research is to pinpoint the concrete mechanisms through which organizations enforce nonviolent discipline in protest.

As an experimental design is unfeasible with this type of research question, the question of causality must be evaluated on the basis of control variables. If there is no omitted variable bias and all relevant confounders are controlled for, the regression has a causal interpretation (Angrist and Pischke 2009). Returning to the control variables used to estimate the models in table 4, it is clearly problematic that the structural variables are not measured on the same level of analysis as the independent and dependent variables. The lack of statistical strength on the structural controls is not surprising as there is little variation on these variables, both across time and space. The surge in temporally and spatially disaggregated research designs point to a need for similarly disaggregated data on structural variables within a polity, such as GDP per capita and unemployment rates. These concerns are to a certain degree alleviated in the autoregressive and fixed effects models, due to the reduction of potential omitted variable bias in these models as discussed above.

Other protest-specific variables that I have not discussed in this thesis may, however, bias the results. This concerns omitted variables that change between week$_{t-1}$ and week$_t$, in particular. If this is the case, causal inferences cannot be drawn from the regression model and the relationships between repression, organization and violent escalation may be spurious. For example, some issues that people take to streets for may of course be more flammable than others, both directly and indirectly. Anti-regime protests may spur a repressive response from the government more quickly than protests with smaller demands, as the former poses a greater threat to the government. Violent escalation of protest may also result from issues that are deeply emotional, and as such trigger an emotional response that is not taken into account by the RA model. This is well-exemplified by the #FMF protests. In isolation, it may be surprising that the protests over fees escalated violently and that police forces were sent to disperse the protest, but considering the deeper causes of structural racism that the protests targeted, the violent escalation is less surprising. The fact that the analysis above does not take into account the issue around which protesters mobilized illustrates the trade-off of analyzing the total protest activity on the ground rather than the campaign. Nonetheless, by treating organization as a variable, the analysis has provided valuable insight into the dynamics of disorganized protest; a phenomenon that is ripe for further research.

A second aspect of protest that this thesis does not address is its magnitude, in terms of number of persons that have mobilized. Well-organized protests may be more able to mobilize additional protesters and thereby appeal to broader segments of the population. This has found to be the comparative advantage of nonviolent protests (Chenoweth and Stephan 2011). The findings above also suggest that organization is important to maintain nonviolent discipline. Large demonstrations may also pose a greater threat to the government, particularly if the protesters are calling for regime change, which may increase the probability of repressive countermeasures from the state. This interaction between protest size and repression is one avenue for further research that can shed light on governments' responses to repression.

However, the inclusion of control variables in the analysis above did not influence the substantive interpretation of the results. This does not eliminate the risk of omitted variable bias. Yet, the fact that the relationships hold when controlling both for structural attributes; protest-specific variables (i.e. duration and spatial diffusion); all omitted variables prior to week$_{t-1}$; and country-year and grid-cell fixed effects increases the likelihood that the models above do have a causal interpretation. From this, three inferences can be drawn: (1) Repression of nonviolent protest leads to violent escalation in the short term if the nonviolent protests are disorganized; (2) Repression of violent protest leads to de-escalation of violent protest in the subsequent week; and (3) Organization leads to violent de-escalation, even in the face of repression.

6 Conclusion

The conclusion of this thesis is written in the midst of a global wave of Black Lives Matter (BLM) demonstrations. In the United States in particular, one of the things that distinguishes the 2020 BLM protests from those of previous years is the level of government repression that the nonviolent protesters are facing. The extensive use of violent techniques of crowd dispersal, alongside the President's threat of evoking the Insurrection Act and dispatching the National Guard to contain the protests (Brooks 2020), exemplifies a situation that closely resembles the research question of this thesis: *Why do some protests escalate violently while others do not?*

The BLM demonstrations started out, and have thus far remained, largely nonviolent, despite some violent encounters between the protesters and police forces. This way, the movement has been able to mobilize supporters from a broad segment of the US population and across country borders. For the protests to remain nonviolent, the results presented in thesis point to one crucial factor: *organization.* I have shown that government repression of nonviolent protest is associated with a shift from nonviolence to violence in the following week. As repression is assumed to be costly, applying the RA model leads to the hypothesis that protesters substitute their chosen tactic - either nonviolent or violent protest - for the other, in the face of repression (Lichbach 1987). The results from this thesis provide additional empirical support for the so-called substitution hypothesis. The large sample used, that not only includes organized protest campaigns, strengthens the external validity of the hypothesis.

Organization seems to be the critical ingredient that helps maintain nonviolent discipline, even when the costs of protesting nonviolently are increasing in the face of repression. This is because partaking in an organized protest increases the individual's cost of resorting to violence, compared to in an unorganized protest (Pinckney 2016). While often assumed, this thesis has empirically shown that organized nonviolent protests are associated with a decrease of violent protests relative to the previous week. The results further show that the association between mild repression of nonviolent protest and violent escalation decreases the more protests are organized. In situations where a high proportion of the protests are organized, the positive association between repression and violent escalation is statistically insignificant. The results suggest that repression is less effective against organized protests, as they can more effectively withstand the costs incurred by repression and thereby maintain and enforce nonviolent discipline. Similar conclusions are drawn regarding repression of violent protest, which is more strongly associated with violent de-escalation where more nonviolent protests are organized.

For the BLM movement to withstand the repressive measures of the government and avoid violent escalation that could have severe consequences and alienate its supporters, organizational capacity “on the ground” is key. Although the protest movement is nationwide, and even global, the demonstrations are local. This thesis stands out from the existing literature by not adopting the campaign as the central unit of analysis. Rather, I have modeled protest events at the time and place where they occurred, using georeferenced protest data from Africa and Asia (Raleigh et al. 2010). Using the PRIO-GRID cell as the unit of analysis has allowed to analyze weekly changes to the relative level of violent protest.

The central explanatory factors, repression and organization, represent characteristics of the two core actors that are active during a protest, namely the government and the protesters, and their interaction. These explanatory factors are modeled with close temporal and spatial proximity to the outcome of interest, which is violent escalation of protest. This is a level of disaggregation seldom applied in the existing literature that addresses protest dynamics, allowing for the analysis of short-term tactical changes.

Another interesting implication from the results presented above is that the concept of violent *escalation* is questionable. Consistently across models, the coefficient of the proportion of violent events in week$_{t-1}$ is negative. In the autoregressive models, the lagged dependent variable also has a consistent negative sign. In other words, high levels of violence in the preceding week is associated with decreasing levels of violence in the current week, controlled for repression and level of organization, as is violent escalation in week$_{t-1}$. This implies that protesters may pick and choose protest tactics, in line with Asal's (2013) conceptualization of an à-la carte menu, where the choice of violence at some point in time does not necessarily imply unidirectional movement along an escalation scale. In fact, the use of violence in week$_{t-1}$ is consistently associated with nonviolence in week$_t$.

The RA model presented in this thesis is parsimoneous in three respects. First, the protesters choices are limited to nonviolent and violent demonstrations. One avenue for further research is to incorporate additional strategic choices within the repertoires of nonviolent *and* violent action. This could include everything from strikes and sit-ins to insurgency and war. Second, the government's choices includes only passivity or repression. Critically, concessions are not considered in this thesis due to data availability. As consistent government responses (i.e. not a mix of concessions and repression) has been argued to be important (Lichbach 1987), the role that concessions play during times of high protest activity is an interesting way forward to fully grasp the dynamics of protest. Furthermore, the consequences of repression has only been analyzed in terms of the choice to protest nonviolently or violently. Other consequences for the regime, such as military and police defection (McLauchlin 2010), or for the protesters, such as increased public support, are gateways for further research. This is particularly relevant in a time where repressive acts are filmed and shared via social media to millions of people almost instantaneously. Finally, the model does not take into account the onset or outcome of demonstrations. While I contend that modeling dynamics during protest is valuable in its own right, further research is needed to assess the consequences of repression and violent escalation on protest outcomes.

In addition to the substantive results, this thesis has demonstrated the shortcomings of structural theories of political violence in the context of short-term events, such as protest. I have shown that structural variables measured on the country-level in annual increments have little explanatory power when assessing the dynamic short-term changes observed *during* protest. While still valuable, the development toward more disaggregated theory and models is necessarily coupled with a need for available disaggregated data. On a final note, methodological advances can provide further insight into whether the results from this thesis are useful for predicting tactical shifts during demonstrations (Ward, Greenhill, and Bakke 2010). After a decade of protest, the time is ripe for further scholarly work on the dynamics of demonstrations.

References

Abrahms, Max. 2006. “Why Terrorism Does Not Work.” *International Security* 31 (2): 42–78.

Addison, Michael. 2002. *Violent Politics: Strategies of Internal Conflict.* New York: Palgrave.

African Economic Outlook 2017. 2017. African Development Bank; OECD; United Nations Development Programme. https://doi.org/10.1787/aeo-2017-en.

Angrist, Joshua D., and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics.* Princeton: Princeton University Press.

———. 2010. “The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con out of Econometrics.” *The Journal of Economic Perspectives* 24 (2): 3–30.

Armed Conflict & Event Data Project (ACLED) Codebook, 2019. 2019. ACLED. https://acleddata.com/acleddatanew/wp-content/uploads/dlm_uploads/2019/04/ACLED_Codebook_2019FINAL_pbl.pdf.

Asal, Victor, Richard Legault, Ora Szekely, and Jonathan Wilkenfeld. 2013. “Gender ideologies and forms of contentious mobilization in the Middle East.” *Journal of Peace Research* 50 (3): 305–18. https://doi.org/10.1177/0022343313476528.

Bell, Curtis. 2016. “The Rulers, Elections, and Irregular Governance Dataset (REIGN).” Broomfield, CO: OEF Research. https://www.oefresearch.org.

Bell, Sam R., and Amanda Murdie. 2016. “The apparatus for violence: Repression, violent protest, and civil war in a cross-national framework.” *Conflict Management and Peace Science* 35 (4): 336–54. https://doi.org/10.1177/0738894215626848.

Bond, Doug, J. Craig Jenkins, Charles L. Taylor, and Kurt Schock. 1997. “Mapping Mass Political Conflict and Civil Society.” *Journal of Conflict Resolution* 41 (4): 553–79.

Bond, Douglas G. 1988. “The Nature and Meanings of Nonviolent Direct Action: An Exploratory Study.” *Journal of Peace Research* 25 (1): 81–89.

Brambor, Thomas, William Roberts Clark, and Matt Golder. 2006. "Understanding Interaction Models: Improving Empirical Analyses." *Political Analysis* 14 (1): 63–82.

Brooks, Rosa. 2020. "Trump wants to crush Black Lives Matter with a law that fought segregation." The Washington Post. June 2020. https://www.washingtonpost.com/outlook/2020/06/02/trump-military-insurrection-act/.

Cameron, A. Colin, and Douglas L. Miller. 2015. "A Practitioner's Guide to Cluster-Robust Inference." *Journal of Human Resources* 50 (2): 317–72.

Cederman, Lars-Erik, Kristian Skrede Gleditsch, and Halvard Buhaug. 2013. *Inequality, Grievances, and Civil War.* Cambridge: Cambridge University Press.

Chenoweth, Erica, Evan Perkoski, and Sooyeon Kang. 2017. "State Repression and Nonviolent Resistance." *Journal of Conflict Resolution* 61 (9): 1950–69.

Chenoweth, Erica, and Maria J. Stephan. 2010. "Mobilization and Resistance: A Framework for Analysis." In *Rethinking Violence: States and Non-State Actors in Conflict*, edited by Erica Chenoweth and Adria Lawrence, 249–76. The MIT Press.

———. 2011. *Why Civil Resistance Works: The Strategic Logic of Nonviolent Conflict.* New York: Columbia University Press.

Christophersen, Knut-Andreas. 2013. *Introduksjon til Statistisk Analyse.* Oslo: Gyldendal Akademisk.

Collier, Paul, and Anke Hoeffler. 2004. "Greed and Grievance in Civil War." *Oxford Economic Papers* 56 (4): 563–95.

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Jan Teorell, David Altman, Michael Bernhard, et al. 2019. "V-Dem Country-Year Dataset 2019." Varieties of Democracy (V-Dem) Project. https://doi.org/10.23696/VDEMCY19.

Cunningham, Kathleen Gallagher. 2013. "Understanding strategic choice: The determinants of civil war and nonviolent campaign in self-determination disputes." *Journal of Peace Research* 50 (3): 291–304. https://doi.org/10.1177/0022343313475467.

Cunningham, Kathleen Gallagher, and Emily Beaulieu. 2010. “Dissent, Repression, and Inconsistency.” In *Rethinking Violence*, edited by Erica Chenoweth and Adria Lawrence, 173–95. Cambridge, Massacusetts: MIT Press.

Cunningham, Kathleen Gallagher, Marianne Dahl, and Anne Frugé. 2017. “Strategies of Resistance: Diversification and Diffusion.” *American Journal of Political Science* 61 (3): 591–605. https://doi.org/10.1111/ajps.12304.

Dahlum, Sirianne, and Tore Wig. 2017. “Educating Demonstrators: Education and Mass Protest in Africa.” *Journal of Conflict Resolution* 63 (1): 3–30. https://doi.org/10.1177/0022002717721394.

———. 2019. “Chaos on Campus: Universities and Mass Political Protest.” *SSRN Electronic Journal*, March. https://doi.org/10.2139/ssrn.3356988.

Davenport, Christian. 2007. “State Repression and Political Order.” *Annual Review of Political Science* 10 (1): 1–23. https://doi.org/10.1146/annurev.polisci.10.101405.143216.

Davenport, Christian, Håvard Mokleiv Nygård, Hanne Fjelde, and David Armstrong. 2019. “The Consequences of Contention: Understanding the Aftereffects of Political Conflict and Violence.” *Annual Review of Political Science* 22 (1): 361–77. https://doi.org/10.1146/annurev-polisci-050317-064057.

Dentlinger, Lindsay. 2018. “#Feesmustfall Damage Costs Soar to Nearly R800M.” Eyewitness News. August 2018. https://ewn.co.za/2018/08/08/feesmustfall-damage-costs-soar-to-nearly-r800m.

Eckstein, Harry. 1980. “Theoretical Approaches to Explaining Collective Political Violence.” In *Handbook of Political Conflict*, edited by Ted Robert Gurr, 135–66. New York: The Free Press.

eNCA. 2015. “University fees in spotlight as Wits protests continue.” eNews Channel Africa. October 2015. https://www.enca.com/south-africa/university-fees-spotlight-during-wits-university-protests.

Esarey, Justin, and Jane Lawrence Sumner. 2017. “Marginal Effects in Interaction Models: Determining and Controlling the False Positive Rate.” *Comparative*

Political Studies 51 (9): 1144–76. https://doi.org/10.1177/0010414017730080.

Essop, Rahima, and Gia Nicolaides. 2015. "#Feesmustfall: Minister Confirms Fee Hikes Won't Exceed 6 Percent." Eyewitness News. October 2015. https://ewn.co.za/2015/10/20/FeesMustFall-Consensus-reached-on-increases.

Eyewitness News. 2015. "STUDENTS VOW TO REMAIN STRONG AFTER COPS FIRE TEAR GAS, STUN GRENADES AT UCT." October 2015. https://ewn.co.za/2015/10/20/Police-fire-teargas-stun-grenades-at-UCT-students-1.

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

Gerring, John. 2005. "Causation." *Journal of Theoretical Politics* 17 (2): 163–98. https://doi.org/10.1177/0951629805050859.

Gleditsch, Kristian Skrede, and Andrea Ruggeri. 2010. "Political opportunity structures, democracy, and civil war." *Journal of Peace Research* 47 (3): 299–310. https://doi.org/10.1177/0022343310362293.

Gustafson, Daniel. 2019. "Hunger to Violence: Explaining the Violent Escalation of Nonviolent Demonstrations." *Journal of Conflict Resolution*, December, 002200271989066. https://doi.org/10.1177/0022002719890669.

Hanck, Christoph, Martin Arnold, Alexander Germer, and Martin Schmelzer. 2019. *Introduction to Econometrics with R.* https://www.econometrics-with-r.org/.

Harris, Adam S., and Erin Hern. 2018. "Taking to the Streets: Protest as an Expression of Political Preference in Africa." *Comparative Political Studies* 52 (8): 1169–99. https://doi.org/10.1177/0010414018806540.

Hegre, Håvard, Tanja Ellingsen, Scott Gates, and Nils Petter Gleditsch. 2001. "Toward a Democratic Civil Peace? Democracy, Political Change, and Civil War, 1816-1992." *American Political Science Review* 95 (1): 33–48.

Hegre, Håvard, Nils W. Metternich, Håvard Mokleiv Nygård, and Julian Wucherpfennig. 2017. "Introduction: Forecasting in Peace Research." *Journal of Peace Research* 54 (2): 113–24. https://doi.org/10.1177/0022343317691330.

Johnston, Richard. 2008. "Survey Methodology." In *The Oxford Handbook of Political Methodology*, edited by Janet M. Box-Steffensmeier, Henry E. Brady, and David Collier, 385–403. Oxford: Oxford University Press.

Kaphahn, Leonie, and Louis Brennan. 2017. "Reassessing the risk conditions for political instability in the light of the Arab Spring." *Journal of Risk Research* 20 (11): 1497–1515. https://doi.org/10.1080/13669877.2016.1179211.

Lichbach, Mark Irving. 1987. "Deterrence or Escalation? The Puzzle of Aggregate Studies of Repression and Dissent." *Journal of Conflict Resolution* 31 (2): 266–97.

Lund, Thorleif. 2002. "Metodologiske prinsipper og referanserammer." In *Innføring i forskningsmetodologi*, edited by Thorleif Lund, 79–123. Oslo: Unipub.

McAdam, Doug, Sidney G. Tarrow, and Charles Tilly. 2011. *Dynamics of Contention.* Cambridge: Cambridge University Press.

McLauchlin, Theodore. 2010. "Loyalty Strategies and Military Defection in Rebellion." *Comparative Politics* 42 (3): 333–50. https://doi.org/10.5129/001041510x12911363509792.

Moore, Will H. 1998. "Repression and Dissent: Substitution, Context, and Timing." *American Journal of Political Science* 42 (3): 851–73.

Mummolo, Jonathan, and Erik Peterson. 2018. "Improving the Interpretation of Fixed Effects Regression Results." *Political Science Research and Methods* 6 (4): 829–35. https://doi.org/10.1017/psrm.2017.44.

Ndlovu, Musawenkosi W. 2017. *#FeesMustFall and Youth Mobilisation in South Africa.* London: Routledge. https://doi.org/10.4324/9781315183435.

Ngcobo, Ziyanda. 2015. "WITS MEDICAL SCHOOL ENTRANCES BLOCKED BY PROTESTERS." Eyewitness News. October 2015. https://ewn.co.za/2015/10/20/Wits-students-blockade-Wits-Medical-School-entrance.

Nicolson, Greg. 2016. "#FeesMustFall: Another day of violence as the state kicks issues forward." Daily Maverick. October 2016. https://www.dailymaverick.co.za/article/2016-10-12-feesmustfall-another-day-of-violence-as-the-state-kicks-issues-forward/#gsc.tab=0.

Opp, Karl-Dieter, and Wolfgang Roehl. 1990. “Repression, Micromobilization, and Political Protest.” *Social Forces* 69 (2): 521–47.

Pearlman, Wendy. 2010. “A Composite-Actor Approach to Conflict Behavior.” In *Rethinking Violence*, edited by Erica Chenoweth and Adria Lawrence, 197–219. Cambridge, Massacusetts: MIT Press.

———. 2012. “Precluding Nonviolence, Propelling Violence: The Effect of Internal Fragmentation on Movement Protest.” *Studies in Comparative International Development* 47 (1): 23–46. https://doi.org/10.1007/s12116-012-9099-2.

Pearson, Michael, Faith Karimi, and David McKenzie. 2015. “South African fee hike off for 2016 after student protests.” CNN. October 2015. https://edition.cnn.com/2015/10/23/africa/south-africa-fee-increase-protests/index.html.

“Per Capita GDP at constant 2015 prices in US Dollars.” 2020. United Nations Statistics Division. March 2020. https://unstats.un.org/unsd/snaama/Downloads.

Pevehouse, Jon C., and Jason D. Brozek. 2008. “Time-series Analysis.” In *The Oxford Handbook of Political Methodology*, edited by Janet M. Box-Steffensmeier, Henry E. Brady, and David Collier, 475–93. Oxford: Oxford University Press.

Pierskalla, Jan Henryk. 2010. “Protest, Deterrence, and Escalation: The Strategic Calculus of Government Repression.” *Journal of Conflict Resolution* 51 (1): 117–45.

Pinckney, Jonathan. 2016. *Making or Breaking Nonviolent Discipline in Civil Resistance Movements*. Washington, DC: ICNC Press.

Poggi, Giovanni. 2015. “Student protests in South Africa have pitted reform against revolution.” The Conversation. November 2015. https://theconversation.com/student-protests-in-south-africa-have-pitted-reform-against-revolution-50604.

Porta, Donatella della. 2012. “Social Movement Studies and Political Violence.” In *The Ashgate Research Companion to Political Violence*, edited by Marie Breen-Smyth, 243–60. Surrey: Ashgate.

Raleigh, Clionadh, Andrew Linke, Håvard Hegre, and Joakim Karlsen. 2010. “Introducing ACLED-Armed Conflict Location and Event Data.” *Journal of Peace Research* 47 (5): 651–60.

Rasler, Karen. 1996. “Concessions, Repression, and Political Protest in the Iranian Revolution.” *American Sociological Review* 61 (1): 132–52.

Riker, William H. 1995. “The Political Psychology of Rational Choice Theory.” *Political Psychology* 16 (1): 23–44.

Roberts, Elizabeth. 2016. “Violence erupts at South Africa student protest.” CNN. October 2016. https://edition.cnn.com/2016/10/04/africa/south-africa-student-protests/index.html.

Rucht, Dieter. 2007. “The Spread of Protest Politics.” In *The Oxford Handbook of Political Behavior*, edited by Russell J. Dalton and Hans-Dieter Klingemann, 708–23. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199270125.003.0038.

Schock, Kurt. 2003. “Nonviolent Action and Its Misconceptions: Insights for Social Scientists.” *PS: Political Science and Politics* 36 (4): 705–12.

———. 2013. “The Practice and Study of Civil Resistance.” *Journal of Peace Research* 50 (3): 277–90. https://doi.org/10.1177/0022343313476530.

———. 2015. “Introduction: Civil Resistance in Comparative Perspective.” In *Civil Resistance: Comparative Perspectives on Nonviolent Struggle*, edited by Kurt Schock, 43:227–57. Social Movements, Protest, and Contentions. Minnesota: University of Minnesota Press.

Sharp, Gene. 1973. *The Politics of Nonviolent Action.* Boston, Massachusetts: Extending Horizons Books.

Shellman, Stephen. 2006. “Process Matters: Conflict and Cooperation in Sequential Government-Dissident Interactions.” *Security Studies* 15 (4): 563–99. https://doi.org/10.1080/09636410601184603.

Shellman, Stephen M., Brian P. Levey, and Joseph K. Young. 2013. “Shifting Sands: Explaining and Predicting Phase Shifts by Dissident Organizations.” *Journal of Peace Research* 50 (3): 319–36. https://doi.org/10.1177/0022343312474013.

Simpson, Brent, Robb Willer, and Matthew Feinberg. 2018. “Does Violent Protest Backfire? Testing a Theory of Public Reactions to Activist Violence.” *Socius: Soci-*

ological Research for a Dynamic World 4 (January): 1–14. https://doi.org/10.1177/2378023118803189.

Stock, James H., and Mark W. Watson. 2007. *Introduction to Econometrics.* Second edition. Boston, Massachusetts: Pearson Addison Wesley.

Stockemer, Daniel. 2014. "What Drives Unconventional Political Participation? A Two Level Study." *The Social Science Journal* 51: 201–11.

Sutton, Jonathan, Charles R. Butcher, and Isak Svensson. 2014. "Explaining Political Jiu-Jitsu: Institution-Building and the Outcomes of Regime Violence Against Unarmed Protests." *Journal of Peace Research* 51 (5): 559–73.

Tarrow, Sidney. 2015. "Contentious Politics." In *The Oxford Handbook of Social Movements*, edited by Donatella Della Porta and Mario Diani. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199678402.013.8.

Tilly, Charles. 2003. *The Politics of Collective Violence.* Cambridge University Press.

Tollefsen, Andreas Forø, Håvard Strand, and Halvard Buhaug. 2012. "PRIO-GRID: A unified spatial data structure." *Journal of Peace Research* 49 (2): 363–74. https://doi.org/10.1177/0022343311431287.

"Unemployment, total (% of total labor force) (modeled ILO estimate)." 2020. International Labour Organization; World Bank. March 2020. https://data.worldbank.org/indicator/sl.uem.totl.zs.

Vestby, Jonas, Vilde Bergstad Larsen, Peder Landsverk, and Andreas Forø Tollefsen. 2020. "PRIO-GRID V3.beta." http:://github.com/prio-data/prio-grid/tree/main.

Ward, Michael D., Brian D. Greenhill, and Kristin M. Bakke. 2010. "The perils of policy by p-value: Predicting civil conflicts." *Journal of Peace Research* 47 (4): 363–75. https://doi.org/10.1177/0022343309356491.

Weidmann, Nils B. 2014. "On the Accuracy of Media-Based Conflict Event Data." *Journal of Conflict Resolution* 59 (6): 1129–49. https://doi.org/10.1177/0022002714530431.

———. 2015. "A Closer Look at Reporting Bias in Conflict Event Data." *American Journal of Political Science* 60 (1): 206–18. https://doi.org/10.1111/ajps.12196.

Weidmann, Nils B., and Michael D. Ward. 2010. "Predicting Conflict in Space and Time." *Journal of Conflict Resolution* 54 (6): 883–901. https://doi.org/10.1177/0022002710371669.

Weiss, D. J., A. Nelson, H. S. Gibson, W. Temperley, S. Peedell, A. Lieber, M. Hancher, et al. 2018. "A Global Map of Travel Time to Cities to Assess Inequalities in Accessibility in 2015." *Nature* 553 (7688): 333–36. https://doi.org/10.1038/nature25181.

Wilkinson, Steven I. 2009. "Riots." *Annual Review of Political Science* 12 (1): 329–43. https://doi.org/10.1146/annurev.polisci.12.041307.075517.

Worrall, John L. 2010. "A User-Friendly Introduction to Panel Data Modeling." *Journal of Criminal Justice Education* 21 (2): 182–96. https://doi.org/10.1080/10511251003693702.

Younge, Gary. 2019. "Streets on Fire: How a Decade of Protest Shaped the World." The Guardian. November 2019. https://www.theguardian.com/culture/2019/nov/23/decade-of-protest-occupy-wall-street-extinction-rebellion-gary-younge.

Østby, Gudrun. 2013. "Inequality and political violence: A review of the literature." *International Area Studies Review* 16 (2): 206–31. https://doi.org/10.1177/2233865913490937.

Appendix

Country coverage

Table 1 shows the countries included in the sample.

Table 1: Countries included in sample

Asia	Africa
Bangladesh	Algeria
Cambodia	Angola
Myanmar	Benin
Nepal	Botswana
Pakistan	Burkina Faso
Sri Lanka	Burundi
Thailand	Cameroon
Vietnam	Central African Republic
	Chad
	Democratic Republic of Congo
	Egypt
	eSwatini
	Ethiopia
	Gabon
	Gambia
	Ghana
	Guinea
	Guinea-Bissau
	Ivory Coast
	Kenya
	Lesotho
	Liberia
	Libya
	Madagascar
	Malawi
	Mali
	Mauritania
	Morocco
	Mozambique
	Namibia
	Niger
	Nigeria
	Republic of Congo
	Rwanda
	Senegal
	Sierra Leone
	Somalia
	South Africa
	South Sudan
	Sudan
	Tanzania
	Togo
	Tunisia
	Uganda
	Zambia
	Zimbabwe

Robustness checks

In this section, I present the alternative model specifications used to assert the robustness of the results presented in the thesis.

Serial correlation

Table 2 shows the base OLS model and the regression of the lagged residual on the residual from the base model. As the coefficient of the lagged residual is statistically signficiant, there is evidence of temporal serial correlation in the data (Worrall 2010, 188). This motivates the use of clustered standard errors, the autoregressive model and fixed effects.

Table 2: Regressing the lagged residual on the residual, base OLS model

	Dependent variable:	
	Change in violent events in week$_t$	Residual
	(1)	(2)
Constant	0.231***	-0.027^{***}
	(0.003)	(0.004)
Residual (lagged)		-0.066^{***}
		(0.015)
Interv. NVP in week$_{t-1}$	0.108***	0.013
	(0.012)	(0.016)
Ex. force NVP in week$_{t-1}$	0.037*	0.001
	(0.022)	(0.027)
Repr. VP in week$_{t-1}$	-0.366^{***}	-0.001
	(0.017)	(0.016)
Prop. organized protest	-0.263^{***}	0.013**
	(0.005)	(0.006)
Prop. VP in week$_{t-1}$	-0.367^{***}	0.058***
	(0.010)	(0.015)
Observations	20,232	14,453
Note:	*p<0.1; **p<0.05; ***p<0.01	

Logistic regression

The logistic regression models are estimated using a binary dependent variable. The model uses a dependent variable where change in the proportion of violent protest is coded 1 where there is an increase, i.e. where $Y_i > 0$ and 0 otherwise. Figure 1 shows the distribution of the logistic dependent variable. Violent escalation - that is, 1's on the dependent variable, comprises 17 percent of the data. Table 3 shows the results of the model.

Positive beta coefficients indicate a positive covariation between the independent variable and the dependent variable. Then, the odds of $Y = 1$ increases with a scale unit increase of X_i. This is the case for intervention against nonviolent protest in both models 4 and 5. The risk of an increase of violent events increases the more nonviolent protests were repressed the previous week. Conversely, negative coefficients, as for the proportion of organized protests, suggest that the odds of $Y_i = 1$ decreases with increasing values of the independent variable. In substantive terms, the odds of violent escalation is lower when a higher number of the protest are organized, in line with H_2 and in line with the results from the OLS regression.

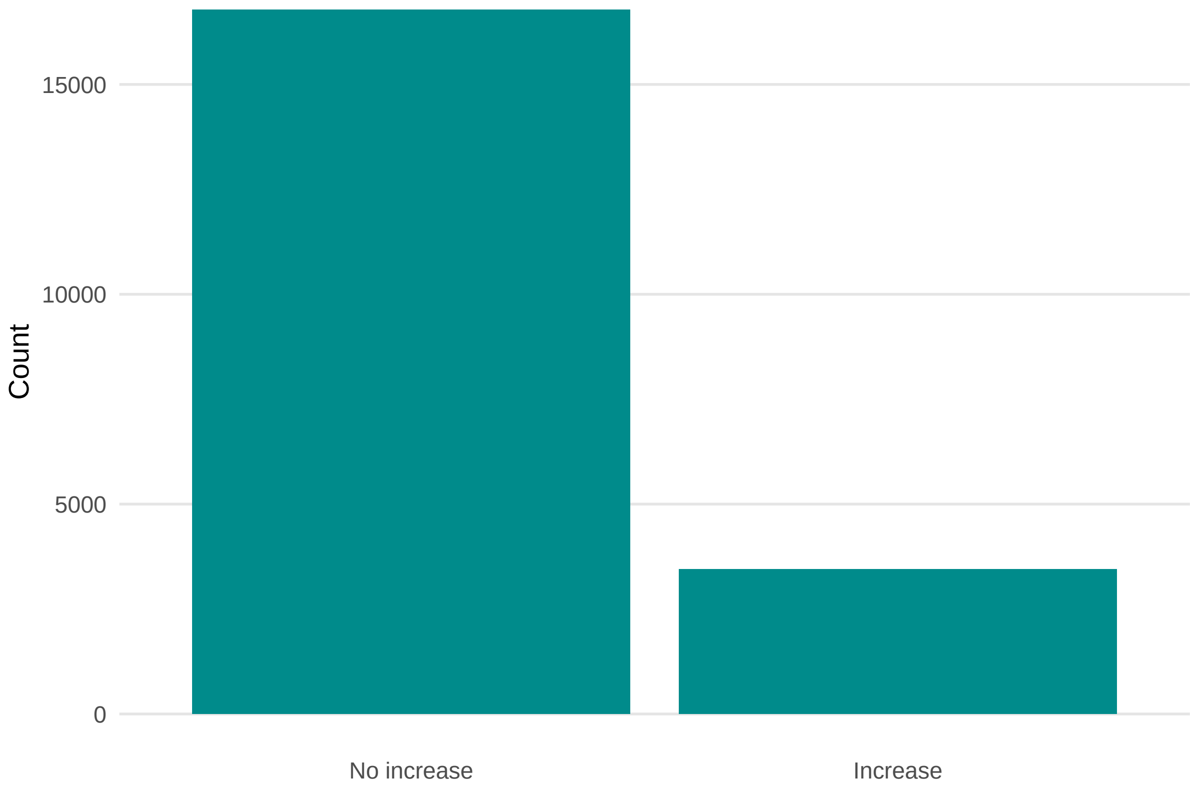

Figure 1: Distribution of the logistic dependent variable

Table 3: Baseline models

	Dependent variable:		
	Increase of violent events in week$_t$		
	Baseline	AR(1)	FE
	(1)	(2)	(3)
Constant	-0.834^{*}	-0.957^{**}	-18.566^{***}
	(0.356)	(0.369)	(0.898)
Interv. NVP in week$_{t-1}$	0.997^{**}	1.218^{***}	0.497^{***}
	(0.331)	(0.343)	(0.102)
Ex. force NVP in week$_{t-1}$	0.119	-0.007	
	(0.311)	(0.322)	
Repr. VP in week$_{t-1}$	-0.976^{***}	-1.354^{***}	-1.445^{***}
	(0.178)	(0.224)	(0.185)
Prop. organized protest	-2.240^{***}	-2.101^{***}	-2.658^{***}
	(0.197)	(0.222)	(0.252)
Prop. VP in week$_{t-1}$	0.223	1.285^{**}	-1.245^{***}
	(0.281)	(0.416)	(0.143)
Y$_{t-1}$		-1.556^{***}	
		(0.189)	
Observations	20,232	14,453	20,232
Log Likelihood	−8,162.145	−5,682.334	−6,513.723
Akaike Inf. Crit.	16,336.290	11,378.670	14,993.450
Fixed effects			Grid-cell
Note:			*p<0.05; **p<0.01; ***p<0.001
			Country-level clustered SEs in parentheses

Full model with excessive force against protesters

Table 4 shows the results from the full model including excessive force against protesters, as well as the interaction term between excessive force against protesters and level of organization. Comparing to Table 3 in Chapter 5.1.1 of the main text, this table shows that the variable is not statistically significant. In addition, its inclusion does not influence the other results, nor is the interaction term significant across any value of the organization variable, visualized in figure 2.

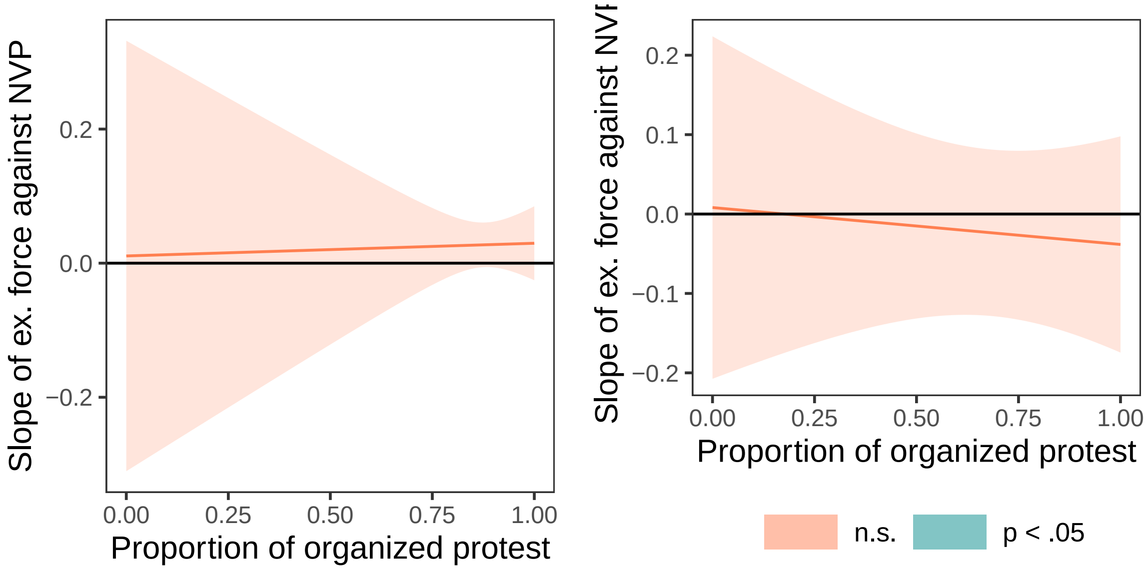

Figure 2: Interaction plots, excessive force and organization, AR(1) and FE

Table 4: Models including excessive force against NVP

	Dependent variable:				
	Change in proportion of violent events in week$_t$				
	Base	AR(1)		FE	
	(1)	(2)	(3)	(4)	(5)
Constant	0.693*** (0.109)	0.197** (0.066)	0.185** (0.065)	−0.000 (0.000)	−0.000
Interv. NVP$^\diamond$	0.068*** (0.014)	0.110*** (0.031)	0.219** (0.077)	0.028* (0.013)	0.123** (0.039)
Ex. force NVP$^\diamond$	0.003 (0.022)	0.027 (0.025)	0.011 (0.079)	−0.016 (0.027)	0.008 (0.053)
Repr. VP$^\diamond$	−0.371*** (0.027)	−0.377*** (0.044)	−0.212** (0.074)	−0.364*** (0.034)	−0.218*** (0.061)
Prop. org. NVP	−0.252*** (0.015)	−0.247** (0.079)	−0.221** (0.075)	−0.245** (0.078)	−0.225** (0.074)
Prop. VP$^\diamond$	−0.412*** (0.026)	−0.191*** (0.055)	−0.189*** (0.055)	−0.556*** (0.045)	−0.553*** (0.045)
Duration (log)	−0.010** (0.003)				
Ongoing VP nb.$^\diamond$	0.002 (0.005)				
Diffusion	0.003 (0.007)				
Avg. travel time	−0.015*** (0.005)				
Lib. dem.	−0.401** (0.125)				
Lib. dem.^{2}	0.624** (0.195)				
GDP/cap. (log)	−0.055*** (0.014)				
Leader tenure	0.00004 (0.0001)				
Elected	0.022** (0.007)				
Unemp. rate	0.007*** (0.002)				
Y$_{t-1}$		−0.211*** (0.024)	−0.210*** (0.022)		
Interv. (NVP) × Prop. Org.			−0.204* (0.087)		−0.191** (0.066)
Exf. (NVP) × Prop. Org.			0.019 (0.090)		−0.046 (0.068)
Repr. (VP) × Prop. Org.			−0.429*** (0.090)		−0.382*** (0.089)
Fixed effects				Grid cell	Grid cell
Observations	20,232	14,453	14,453	20,232	20,232
Adjusted R^2	0.280	0.303	0.313	0.374	0.381
Residual Std. Error	0.287 (df = 20216)	0.277 (df = 14446)	0.275 (df = 14443)	0.268 (df = 19248)	0.267 (df = 19245)
F Statistic	526.395*** (df = 15; 20216)	1,046.381*** (df = 6; 14446)	732.932*** (df = 9; 14443)	13.300*** (df = 983; 19248)	13.630*** (df = 986; 19245)

Note: *p<0.05; **p<0.01; ***p<0.001
Country-level clustered SEs in parentheses
$^\diamond$ Lagged one week

Fixed effects

Table 5 shows various models with fixed effects on the grid-cell, country, and country-year level; with and without the lagged dependent variable. By and large, the results from the models reported in the thesis are robust across these models.

Figures 3 and 4 show the interaction plots between repression and organization from the fixed effects model reported in Table 4 in the thesis. When comparing to the AR(1) plots reported in the thesis, the substantial results are unchanged when incorporating fixed effects.

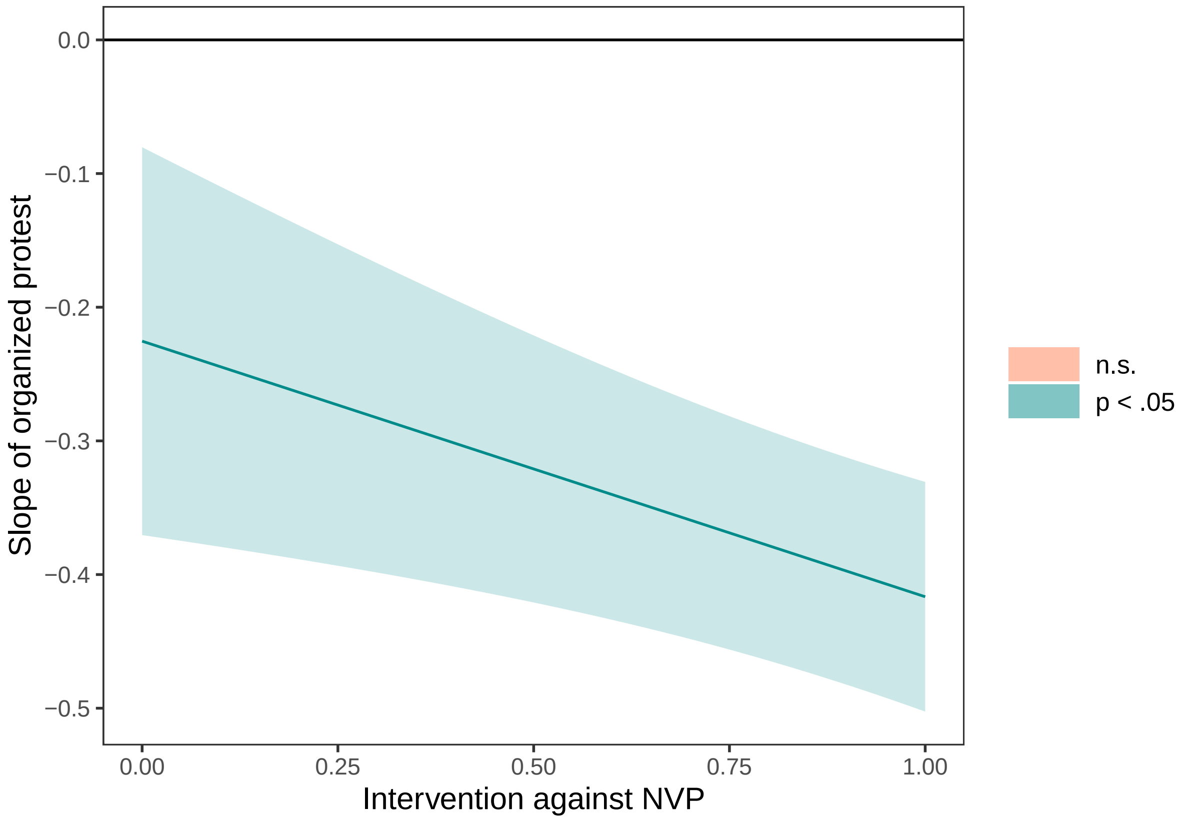

Figure 3: Interaction plot, interv. (NVP) in week$_{t-1}$ and prop. organized protests, FE

Table 5: Fixed effects

	Dependent variable:				
	Change in proportion of violent events in week$_t$				
	AR(1) + FE	FE			
	(1)	(2)	(3)	(4)	(5)
Constant	0.186*** (0.033)	−0.000 (0.000)	0.718*** (0.110)	0.115 (0.235)	0.052 (0.235)
Interv. NVP◇	0.035 (0.020)	0.024 (0.020)	0.065*** (0.014)	0.036** (0.014)	0.034* (0.014)
Repr. VP◇	−0.397*** (0.037)	−0.371*** (0.034)	−0.372*** (0.027)	−0.381*** (0.025)	−0.381*** (0.024)
Prop. org. NVP	−0.244** (0.085)	−0.238** (0.082)	−0.251*** (0.015)	−0.255*** (0.015)	−0.255*** (0.015)
Prop. VP◇	−0.375*** (0.051)	−0.491*** (0.057)	−0.416*** (0.025)	−0.491*** (0.021)	−0.493*** (0.021)
Y_{t-1}	−0.125*** (0.018)	−0.093*** (0.023)			
Duration (log)			−0.009** (0.003)	−0.001 (0.003)	−0.001 (0.003)
Ongoing VP nb.◇			0.001 (0.005)	0.007 (0.005)	0.007 (0.005)
Diffusion			0.003 (0.007)	0.005 (0.007)	0.006 (0.007)
Avg. travel time			−0.015*** (0.004)	0.009* (0.004)	0.009* (0.004)
Lib. dem.			−0.397** (0.125)	−0.171 (0.233)	−0.211 (0.245)
Lib. dem.^{2}			0.631** (0.197)	0.090 (0.314)	0.179 (0.336)
GDP/cap. (log)			−0.056*** (0.014)	0.003 (0.030)	0.006 (0.030)
Leader tenure			0.0001 (0.0001)	−0.0001 (0.0001)	−0.00004 (0.0001)
Elected			0.029** (0.009)	0.005 (0.006)	0.003 (0.007)
Unemp. rate			0.007*** (0.002)	0.006 (0.004)	0.011* (0.005)
Fixed effects	Country	Grid-cell	Year	Country	Country-year
n	54	979	9	54	371
Observations	14,453	14,453	20,232	20,232	20,232
Adjusted R^2	0.356	0.400	0.282	0.328	0.329
Residual Std. Error	0.266 (df = 14397)	0.257 (df = 13901)	0.287 (df = 20209)	0.278 (df = 20164)	0.278 (df = 20156)
F Statistic	146.376*** (df = 55; 14397)	18.490*** (df = 551; 13901)	361.662*** (df = 22; 20209)	148.490*** (df = 67; 20164)	133.271*** (df = 75; 20156)

Note: *p<0.05; **p<0.01; ***p<0.001
Country-level clustered SEs in parentheses
◇ Lagged one week

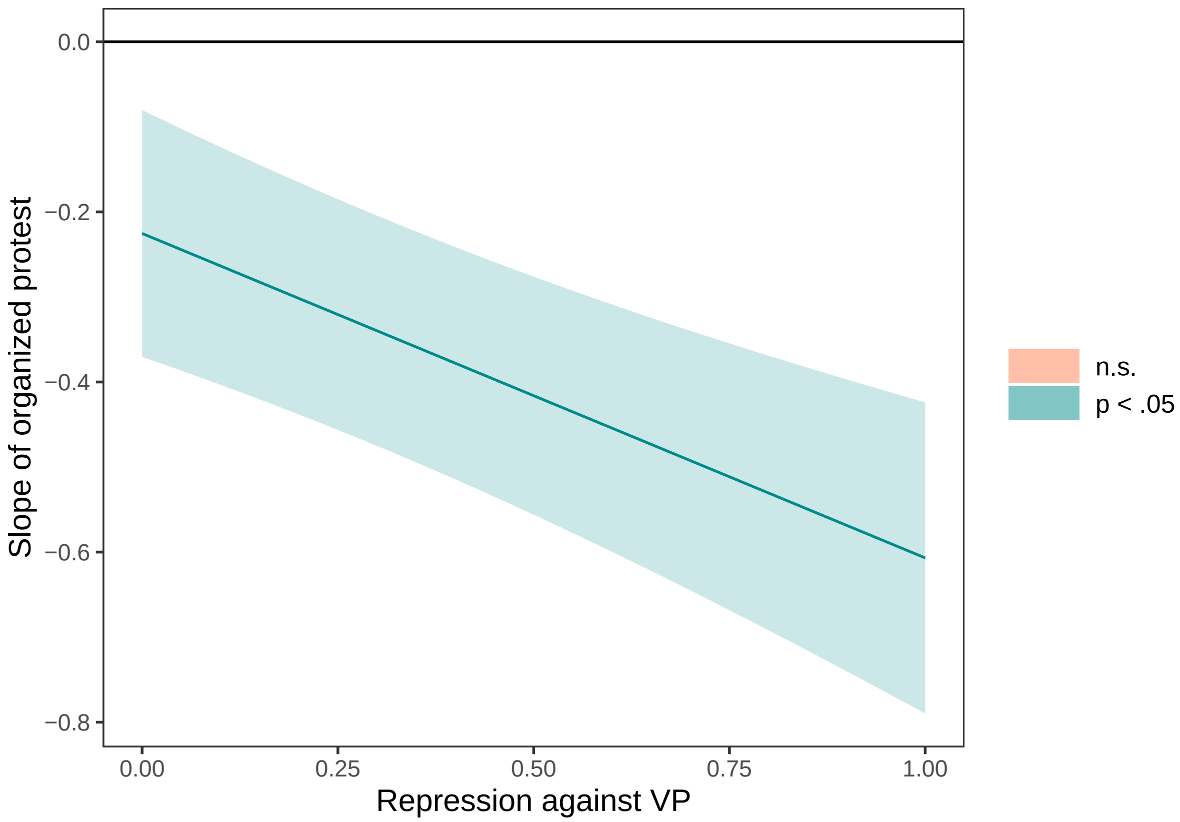

Figure 4: Interaction plot, repr. (VP) in week$_{t-1}$ and prop. organized protests, FE

OLS, binary independent variables

The following subset of models, reported in Table 6, are estimated using dichotomous rather than continuous independent variables. The variables are coded 1 if the original variable is greater than 0, and 0 otherwise. In other words, the variables identify whether repression occurred or did not occur and whether one or more nonviolent protests are organized or not. With this operationalization of the independent variables, we see that the results from the models reported in the thesis are robust in terms of direction of the coefficients. Figures 3 and 4 show the interaction term, visualizing where the coefficients are statistically significant.

Table 6: Dichotomous independent variables

	Dependent variable:			
	Change in proportion of violent events in week$_t$			
	AR(1)	OLS	AR(1)	FE
	(1)	(2)	(3)	(4)
Constant	0.207** (0.063)	0.225** (0.069)	0.201** (0.065)	0.000 (0.000)
Interv. NVP$^\diamond$	0.072*** (0.014)	0.114** (0.041)	0.111* (0.047)	0.074** (0.028)
Repr. VP$^\diamond$	−0.156*** (0.034)	−0.055 (0.054)	−0.114 (0.060)	−0.118* (0.051)
Org. NVP	−0.195*** (0.059)		−0.186** (0.062)	−0.186** (0.062)
Prop. org NVP		−0.243*** (0.072)		
Prop. VP$^\diamond$	−0.229*** (0.062)	−0.418*** (0.046)	−0.232*** (0.062)	−0.598*** (0.048)
Y_{t-1}	−0.226*** (0.026)		−0.226*** (0.025)	
Interv. (NVP)$^\diamond$ × Prop. Org.		−0.090 (0.047)		
Repr. (VP)$^\diamond$ × Prop. Org.		−0.242*** (0.057)		
Interv (NVP)$^\diamond$ × Org. (dummy)			−0.052 (0.050)	−0.063 (0.038)
Repr. (VP)$^\diamond$ × Org. (dummy)			−0.060 (0.041)	−0.046 (0.040)
Observations	14,453	20,232	14,453	20,232
Adjusted R^2	0.273	0.253	0.274	0.344
Fixed effects				Grid cell

Note: *p<0.05; **p<0.01; ***p<0.001
Country-level clustered SEs in parentheses

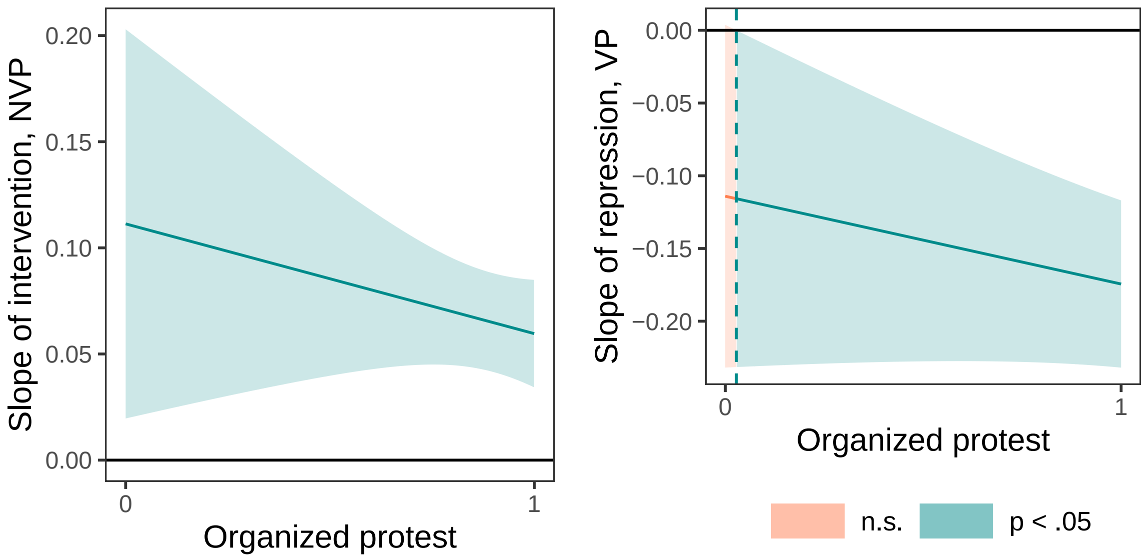

Figure 5: Interaction plots, dummy variables (AR(1) model)

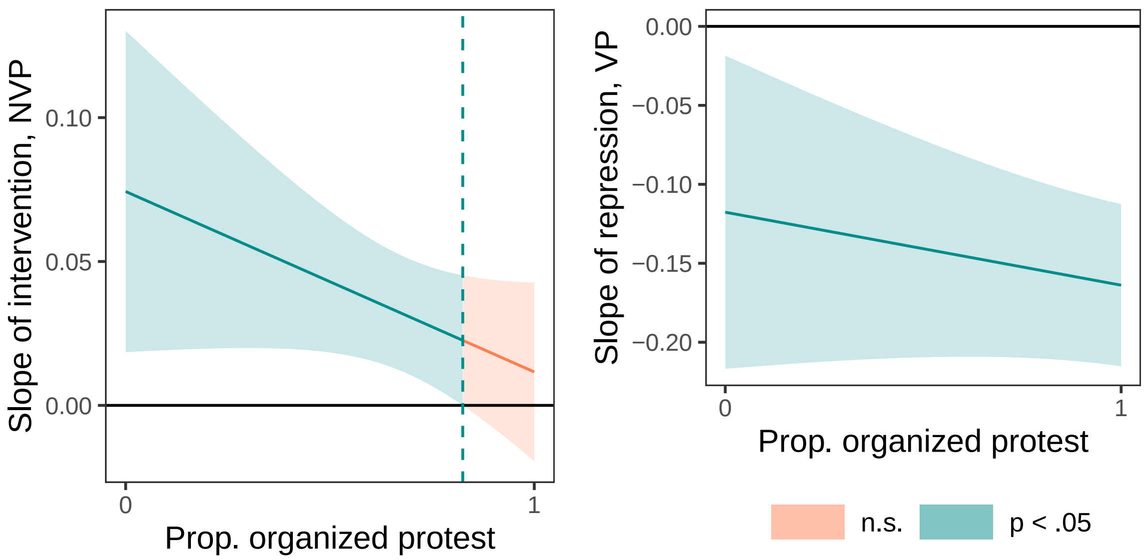

Figure 6: Interaction plots, dummy variables (FE model)

Increase of violent events

In Table 7, the key OLS models are estimated using a dependent variable that only measures increases of violent events in week$_{t-1}$. Decreases, i.e. where the original dependent variable is negative, are set to zero. Still, negative coefficients indicate a decrease of violent events relative to the preceding week, while positive values imply an increase. The results from the models reported in the thesis hold across these models.

Table 7: Increase

	Dependent variable:			
	Increase in proportion of violent events in week$_t$			
	OLS	AR(1)	FE	AR(1)
	(1)	(2)	(3)	(4)
Constant	0.211*** (0.062)	0.175** (0.057)	−0.000	0.174** (0.060)
Interv. NVP$^\diamond$	0.110*** (0.033)	0.113*** (0.029)	0.038*** (0.011)	0.231*** (0.070)
Repr. VP$^\diamond$	−0.138*** (0.029)	−0.145*** (0.030)	−0.133*** (0.023)	−0.211*** (0.053)
Prop. org. NVP	−0.220*** (0.058)	−0.190*** (0.056)	−0.204*** (0.060)	−0.190** (0.062)
Prop. VP$^\diamond$	0.014 (0.032)	0.139*** (0.038)	−0.140*** (0.027)	0.140*** (0.038)
Y_{t-1}		−0.139*** (0.015)		−0.139*** (0.015)
Interv. (NVP) × Prop. Org.				−0.227** (0.076)
Repr. (VP) × Prop. Org.				0.173* (0.074)
Observations	20,232	14,453	20,232	14,453
Adjusted R^2	0.129	0.136	0.268	0.142
Fixed effects			Grid cell	

Note: *p<0.05; **p<0.01; ***p<0.001
Country-level clustered SEs in parentheses
$^\diamond$ Lagged one week

Standard errors

The models in Table 8 report heteroskedasticity-robust standard errors, and Table 9 reports standard errors clustered on the grid cell. As we can see, the standard errors in both tables are much lower than the country-clustered standard errors reported in the thesis.

Table 8: Heteroskedasticity-robust SEs

	Dependent variable:		
	Change in proportion of violent events in week$_t$		
	AR(1)		FE
	(1)	(2)	(3)
Constant	0.198^{***} (0.005)	0.185^{***} (0.005)	-0.000
Interv. NVP$^{\diamond}$	0.110^{***} (0.017)	0.219^{***} (0.038)	0.123^{***} (0.028)
Repr. VP$^{\diamond}$	-0.376^{***} (0.026)	-0.212^{***} (0.034)	-0.218^{***} (0.032)
Prop. org. NVP	-0.247^{***} (0.006)	-0.221^{***} (0.006)	-0.225^{***} (0.005)
Prop. VP$^{\diamond}$	-0.191^{***} (0.022)	-0.190^{***} (0.022)	-0.553^{***} (0.018)
Y_{t-1}	-0.211^{***} (0.014)	-0.210^{***} (0.014)	
Interv. (NVP) × Prop. Org.		-0.204^{***} (0.042)	-0.191^{***} (0.032)
Repr. (VP) × Prop. Org.		-0.429^{***} (0.033)	-0.381^{***} (0.032)
Observations	14,453	14,453	20,232
Adjusted R^2	0.303	0.313	0.381

Note: $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$
Heteroskedasticity-robust SEs in parentheses
$^{\diamond}$ Lagged one week

Table 9: Grid-cell clustered SEs

	Dependent variable:		
	Change in proportion of violent events in week_t		
	AR(1)		FE
	(1)	(2)	(3)
Constant	0.198^{***} (0.016)	0.185^{***} (0.015)	−0.000 (0.000)
Interv. $\text{NVP}^{\diamond}$	0.110^{***} (0.016)	0.219^{***} (0.037)	0.123^{***} (0.028)
Repr. $\text{VP}^{\diamond}$	-0.376^{***} (0.027)	-0.212^{***} (0.041)	-0.218^{***} (0.042)
Prop. org. NVP	-0.247^{***} (0.018)	-0.221^{***} (0.017)	-0.225^{***} (0.015)
Prop. $\text{VP}^{\diamond}$	-0.191^{***} (0.033)	-0.190^{***} (0.032)	-0.553^{***} (0.022)
Y_{t-1}	-0.211^{***} (0.017)	-0.210^{***} (0.016)	
Interv. (NVP) × Prop. Org.		-0.204^{***} (0.042)	-0.191^{***} (0.032)
Repr. (VP) × Prop. Org.		-0.429^{***} (0.048)	-0.381^{***} (0.052)
Observations	14,453	14,453	20,232
Adjusted R^2	0.303	0.313	0.381

Note: $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$
Grid-cell clustered SEs in parentheses
$^{\diamond}$ Lagged one week

Robust regression

Table 10 shows the results from a robust regression of the base autoregressive model. The direction of the coefficients are robust in comparison to the results reported in the thesis, but the values of the coefficients change when using robust regression.

Table 10: Robust regression

	Dependent variable:
	Change in proportion of violent events in week$_t$
Constant	0.007**
	(0.003)
Interv. NVP in week$_{t-1}$	0.005***
	(0.002)
Repr. VP in week$_{t-1}$	$-$0.894***
	(0.225)
Prop. organized NVP	$-$0.009**
	(0.004)
Prop. VP in week$_{t-1}$	$-$0.016
	(0.028)
Y$_{t-1}$	$-$0.033**
	(0.013)
Observations	14,453
Residual Std. Error	0.011 (df = 14447)
Note:	*p<0.1; **p<0.05; ***p<0.01
	Country-level clustered SEs in parentheses

Modeling only increase

Table 11 shows the base models from Chapter 6.1, but observations that have zero or negative values on the dependent variable are filtered out of the sample, introducing a potential selection bias. The results are nonetheless robust.

Table 11: Observations where DV > 0 only

	Dependent variable:			
	Increase in proportion of violent events in week$_t$			
	OLS	AR(1)	FE	AR(1)
	(1)	(2)	(3)	(4)
Constant	0.829***	0.762***	0.952***	0.763***
	(0.028)	(0.041)	(0.008)	(0.041)
Interv. NVP in week$_{t-1}$	0.062**	0.075*	0.018	0.074
	(0.030)	(0.043)	(0.026)	(0.056)
Repr. VP in week$_{t-1}$	−0.818***	−0.865***	−0.718***	−0.884***
	(0.055)	(0.062)	(0.070)	(0.063)
Prop. organized NVP	−0.926***	−0.842***	−0.728***	−0.845***
	(0.018)	(0.020)	(0.050)	(0.022)
Prop. VP in week$_{t-1}$	−0.109***	−0.010	−0.161***	−0.011
	(0.023)	(0.044)	(0.020)	(0.044)
Y$_{t-1}$		−0.009		−0.009
		(0.025)		(0.025)
Interv. NVP × Prop. Org.				0.005
				(0.071)
Repr. VP × Prop. Org				0.184*
				(0.100)
Observations	3,454	2,404	3,454	2,404
Adjusted R^2	0.538	0.490	0.628	0.489
Note:	*p<0.1; **p<0.05; ***p<0.01			
	Country-level clustered SEs in parentheses			

OLS vs. fixed vs. random effects

The results show that fixed effects are preferred over regular OLS and random effects modeling.

```r
acled_panel <- pdata.frame(acled_cleaned, index = c("gid", "week"))


ols <- lm(chg_viol ~ interv_lag_prop +  repviol_lag_prop +
                 prop_organized +
                 prop_viol_lag, data = acled_panel)
fixed <- plm(chg_viol ~ interv_lag_prop +  repviol_lag_prop +
               prop_organized +
               prop_viol_lag, data = acled_panel, model = "within")
print_output(pFtest(fixed, ols))
```

```
 F test for individual effects

data:  chg_viol ~ interv_lag_prop + repviol_lag_prop + prop_organized +  ...
F = 4.9333, df1 = 978, df2 = 19249, p-value < 2.2e-16
alternative hypothesis: significant effects
```

```r
random <- plm(chg_viol ~ interv_lag_prop +  repviol_lag_prop +
                prop_organized +
```

```r
                 prop_viol_lag, data = acled_panel, model = "random")
print_output(phtest(fixed, random))
```

```
	Hausman Test

data:  chg_viol ~ interv_lag_prop + repviol_lag_prop + prop_organized +  ...
chisq = 526, df = 4, p-value < 2.2e-16
alternative hypothesis: one model is inconsistent
```

References

Worrall, John L. 2010. "A User-Friendly Introduction to Panel Data Modeling." *Journal of Criminal Justice Education* 21 (2): 182–96. https://doi.org/10.1080/10511251003693702.