Universität Bayreuth
Rechts- und Wirtschaftswissenschaftliche Fakultät
Lehrstuhl VWL V - Institutionenökonomik
Professor Dr. Martin Leschke

Rational Dissent: Economic Perspectives on Protesting

Dissertation
zur Erlangung des Grades eines Doktor der Wirtschaftswissenschaft der Rechts- und Wirtschaftswissenschaftlichen Fakultät der Universistät Bayreuth

Vorgelegt von
Tim Baule
aus Groß-Umstadt

Dekan: Prof. Dr. Claas Christian Germelmann (Universität Bayreuth)
Vorsitz der Prüfungskommission: Prof. Dr. Amelie Wuppermann (Universität Bayreuth)
Erstberichterstatter: Prof. Dr. Martin Leschke (Universität Bayreuth)
Zweitberichterstatter: Prof. Dr. Sebastian Till Braun (Universität Bayreuth)
Tag der mündlichen Prüfung: 12.11.2025

Acknowledgements

First and foremost, I want to thank Martin Leschke for his tremendous support over the past three years. This support included not only regular feedback on research projects but also a great deal of patience in discussing any topic that came to my mind during the weekly *Oberseminar* at the Chair VWL V. The past three years have been a constant learning experience, teaching me far more about science, economics, and related topics than what is presented in the following pages.

All research articles in this dissertation benefited substantially from the feedback of various fellow researchers. For the article "Resources and Protest Mobilization", I thank the participants of the following research conferences: Annual Conference of the Committee for Institutional Economics of the Verein für Socialpolitik 2024, 6th International Conference on European Economics and Politics, Warsaw International Economic Meeting 2024, Bayreuth Graduate Seminar 2024, Radein Seminar 2024, Bavarian Graduate Program in Economics Research Workshop 2024. Furthermore, I am grateful for helpful suggestions and feedback from Martin Leschke, Jonathan Bothner, Tim Röthel, and Felix Schlieszus.

For the second article "From Individuals to Networks", I thank my co-authors Jonathan Bothner and Maximilain Kähny, as well as Fabian Herweg, Henrik Guhling, Martin Leschke, and conference and seminar participants at the Public Choice Conference 2025, University of Bayreuth Graduate Seminar 2024, Bavarian Micro Day 2024, and Jahrestagung des Ausschusses für Wirtschaftssysteme und Institutionenökonomik 2024 for helpful comments and suggestions. I also thank the Institute for Humane Studies for providing travel support to attend the Public Choice Conference 2025. Additionally, I am honored to have received the Elinor and Vincent Ostrom Prize for the best paper by a graduate student at the Public Choice Conference 2025.

For the third paper "From Silence to Salience", I thank my co-author, Henrik Guhling, as well as the participants of the following research conferences: Public Choice Conference 2025, Radein Seminar 2025, Bayreuth Graduate Seminar 2024 and 2025, and 38th BGPE Research Workshop. I thankfully acknowledge helpful suggestions and feedback from Martin Leschke and Fabian Herweg. Also, I would like to thank RW Alumni Bayreuth and the Institute for Humane Studies for providing travel support.

My learning experience in Bayreuth would not have been the same without my colleagues, with whom I shared my struggles, successes, and three wonderful years of doctoral studies. I am especially grateful to the Skat group (Jonathan Bothner, Alessandro Kadner-Graziano, Maximilian Kähny, Tim Röthel, and Felix Schlieszus) and the later generation of friends with whom I still went to Mensa at 11:30 AM every day (Henrik Guhling, Tim

Lüdiger, Timur Öztürk, and Christoph Schlömer). You all made my time in Bayreuth truly special.

Finally, I would like to thank Ella, my friends, and my family for their unwavering support. I could not have succeeded without you!

Summary

The topic of this cumulative dissertation is protesting as collective action. It begins with a general introduction to the subject, followed by three research articles that contribute to our understanding of individual protest decisions and the overall occurrence of protest events. The dissertation ends with a general conclusion that ties all the elements together.

In the first research article, the main question is how resources affect the occurrence of protest events at the county level in the US. The main hypotheses suggest that traditional organizations, such as NGOs, and online networks formed on social media platforms significantly enhance protest mobilization capacities. The findings indicate that both resource dimensions are associated with an increase in protest incidence; however, they interact negatively, suggesting a substitutive effect on mobilizations. Additionally, resources, especially traditional organizations, are found to mediate grievance-based protesting.

The second research article examines the role of networks in individual protest decisions. A microeconomic model formalizes equilibrium solutions for protest outcomes, which are then tested using a spatial autoregressive model. The findings reveal the existence of network effects across the entire sample, as well as for more homogeneous subgroups within the US population. These results demonstrate the existence of spillover effects in protest participation. Furthermore, they provide a novel explanation for the observed sizes of four major protest movements in the US.

Finally, the third research article proposes a novel attention mechanism for protest coordination. It adapts salience theory and applies it to the field of political economy to formally model how high-profile media events affect individual perceptions of protest events. The model's implications are empirically assessed by examining police killings of Black individuals. The results speak in favor of attention-based protest mobilization and indicate this channel to have substantial explanatory power for short-term mobilization patterns.

Contents

Chapter 1

General Introduction

> “[P]ublic choice may be summarized by the three-word description, ‘politics without romance’.” — *James M. Buchanan Jr., 2003*

Humans have long lived in communities that require organization and cooperation. Politics, in this context, arises from the necessity of enabling many individuals to live and thrive within a shared society. Within such societies, there are various ways to express one’s political demands, and one particular form of political expression is the protagonist of this dissertation: protesting. Throughout the past centuries, protest has consistently proven to be a powerful force for both driving societal transformation and generating conflict. Historical examples abound: the civil rights movement fought for racial equality in the United States; Mahatma Gandhi led the Salt March in defiance of British colonial taxation; and the Anti-Apartheid Movement in South Africa played a pivotal role in establishing a democratic system based on equal voting rights. These cases, among many others, underscore the central role of protest in political life and, by extension, in the broader organization of human societies.

The significance of protesting in shaping political processes motivates the endeavor to understand and explain the conditions under which protests emerge and the factors influencing individual participation. Theoretically, such insights contribute to our understanding of political processes, particularly regarding political expression and communication. Practically, they contribute to understanding contemporary political developments and their drivers. This dissertation addresses three distinct questions in the context of protest emergence: (1) What is the role of different societal resources in protest events? (2) How do social networks affect individual mobilization? (3) What role do attention-grabbing events play in coordinating collective action? To explore these questions, the dissertation applies economic theory and applied econometrics, aiming to deepen our understanding of mobilization dynamics in protest settings. The individual papers are connected by a shared explanandum: the phenomenon of protest and the decision-making processes underlying individual participation.

From a methodological point of view, the dissertation contributes to understanding protest activity by applying tools from the Public Choice research program within the field of political economy. Buchanan's (2003) quote at the outset summarizes the fundamental logic of this approach: to analyze politics without romanticizing. In the spirit of this perspective, the guiding principle of this dissertation could be summarized as "protests without romance." So, protesting should not be understood as a heroic, selfless act of altruistic agents sacrificing their effort, time, and safety to fight for the common good. Instead, protesting should be understood as political actors who, very much like market participants, respond to unique preferences and constraints, seeking to maximize their individual utility. Consequently, no idealized portrayal of protesters or the act of protesting itself is presented in this dissertation. Protest decisions are modeled as the outcomes of rational decision-makers trying to fulfill their competing objectives with limited resources, embedded in specific social and institutional contexts.

The assumption of rational actors should not be understood as an empirical fact but rather as something useful for explanations in the social sciences. According to Popper (1994, p. 164), researchers should see the rationality principle as "animating law" or "approximation of the truth" to understand how individuals react to different situational characteristics. In this program, the rationality principle should be granted, and the researchers should focus on the "situational analysis," which characterizes the circumstances and preferences of individuals (Popper, 1994, p. 166). Popper (1994, p. 178) justifies this methodological choice by arguing that models constructed based on this structure outperform alternative modeling approaches in terms of scientific fruitfulness, scope, and explanatory power.

A strict application of the fundamental economic axioms, i.e., methodological individualism and Homo Oeconomicus, led Olson (1971) to an insightful theory about *The Logic of Collective Action*. As a first insight, protest should be understood as a public good that serves group interests. That is, everyone who shares the objectives of the protest benefits from the signal that the protest event sends to politicians, potentially because it increases the likelihood of policy changes in their favor. However, the contribution to the protest of each individual is marginal; that is, a protest needs to be large to send a strong signal, but one individual more or less does not change the nature of the signal. In this context, the rational decision of each individual is to abstain from protesting, as its benefits can be enjoyed anyway, but the costs of participation can be avoided. As a result, the classical free-rider problem emerges, as no one contributes to the public good of protesting, and we observe zero participation.

Given this mechanism, Olson (1971) highlights the importance of *selective incentives* as a way to enable collective action. Selective incentives are private benefits an individual experiences conditional on contributing to the provision of the public good. Feelings of moral duty or network effects are examples of such selective incentives. For example, suppose an individual is embedded in a social circle that judges it morally praiseworthy to protest and in which social gatherings may occur at protest events. In that case, the private participation cost may be outweighed by the benefits derived from one's social circle. Positive participation in the provision of public goods is thus explained by relying on selective incentives. Importantly, the extent to which the public good is provided does not affect the individual incentive to join the protest, as the individual contribution remains marginal.

Various variables and mechanisms are discussed in the literature that (potentially) enable collectives to overcome the free-rider problem by providing selective incentives. Within this literature, this dissertation explores three specific aspects. These include the endowments of resources that facilitate the organization of protests, the presence of networks within which individuals make their protest decisions, and the impact of attention-grabbing events in the context of protest mobilization. Each paper stands on its own as a contribution to a specific scientific debate, and they stand jointly as a collection of scientific work aimed at understanding protesting as a form of collective action.

The research article presented in Chapter 2, *Resources and Protest Mobilization – Substitution and Mediation in Collective Action*, contributes to a strand of literature that examines the role of resources in enabling and facilitating protest mobilization. In this article, the three resources–societal organizations, online networks, and financial means–are conceptualized and operationalized using principal component analysis. In a subsequent empirical analysis, all resource dimensions demonstrate a positive impact on the incidence of protest events at the US county level between 2018 and 2022. The results indicate a substitutive relationship between societal organizations and online networks in their effects on protest mobilization. Furthermore, grievances are crucially mediated via resources, especially societal organizations. The paper connects to a rich literature in Resource Mobilization Theory by disentangling separate resource dimensions and highlighting their substitutive nature. Furthermore, the paper reconciles grievance-based and resource-based approaches by identifying resources as a crucial mediator of grievances.

Rationalizing Protests – From Individuals to Networks is the research article presented in Chapter 3, investigating the role of networks in individual protest decisions. Motivated by an adapted model of network goods from Industrial Organization, the paper assesses standalone and network benefits in individual protest decisions. Empirically, the analysis applies a spatial autoregressive model, which, unlike standard inference, allows for spillover effects among individuals. The results indicate both the presence of standalone and network benefits and specify which subgroups of the network effects dominate the standalone benefits. The research article connects to a growing body of literature on network effects in protest decisions. Unlike previous work, the analysis is conducted with a representative sample of the national population (US). The application of spatial econometric techniques in the protest context is also novel and allows for distinctly identifying network effects.

The final research article, *From Silence to Salience – How Critical Events Shape Protests*, in Chapter 4, proposes a novel attention mechanism to the protest literature by adapting saliency theory. As a first step, the theoretical model introduces the channel that mediates the effects of events that receive high media attention (critical events) on protest outcomes. Empirically, this channel is tested by relying on data from the Black Lives Matter protests between 2018 and 2022. We interpret police killings of Black citizens as critical events for protest mobilization. Then, we assess the effect of these events on protest mobilization directly and via the attention channel. As a result, we conclude that the modeled attention channel can explain large parts of the short-term mobilization. This paper contributes to the literature by proposing a new formal channel that takes into account the contextual dependencies of individual protest decisions. In the subsequent empirical analysis, this channel is assessed and identified as explaining a large fraction of the subsequent short-term mobilization.

Chapter 2

Resources and Protest Mobilization – Substitution and Mediation in Collective Action

Abstract: This article examines the role of three key resource dimensions–societal organizations, online networks, and financial endowments–in shaping protest mobilization across US counties. Each resource dimension is operationalized relying on principal component analysis and incorporated into a Poisson quasi-maximum likelihood estimation. The results show that all three dimensions significantly increase the expected number of protest events. Notably, societal organizations and online networks show a substitutive relationship in mobilization processes. Moreover, the analysis reveals that resources, particularly societal organizations, mediate the effect of grievances on protest activity. These findings contribute to the literature on Resource Mobilization Theory by disentangling different resource dimensions and assessing their mediating role for grievances.[1]

Keywords: Protests; Demonstrations; Collective Action; Resources; Resource Mobilization Theory

[1]This research article is currently under review at the Journal of Peace Research. The replication code can be accessed at `https://doi.org/10.7910/DVN/CZENVY` and the working paper version is available at `https://doi.org/10.15495/EPub_UBT_00007846`.

2.1 Introduction

Over the past few decades, protest dynamics have undergone substantial transformation, largely driven by innovations in information and communication technologies (ICTs). As described by Earl (2018), the Arab Spring protests beginning in 2010 demonstrate the far-reaching potential of online networks for protest coordination. These digital platforms serve as decentralized tools that enable citizens to overcome traditional barriers to collective action. Historically, societal organizations such as churches, unions, and non-governmental organizations (NGOs) fulfilled this function by providing essential mobilizing resources, including communication infrastructure and organizational capabilities (Edwards and Gillham, 2013). However, the technological innovations offer fundamentally new coordination mechanisms, shifting the importance from hierarchical, centralized organizations to non-hierarchical, decentralized online networks (Shirky, 2011).

This article examines the role of societal organizations, online networks, and financial endowments in protest mobilization at the US county level. The primary objective is to disentangle the former two resource dimensions in order to compare their respective effect sizes and interdependencies. Financial endowments are incorporated as an additional category of mobilizing capacity, as they may confound the effects of the other resource dimensions. In a second step, the analysis investigates the mediating role of resources in grievance-based protesting. The underlying rationale is that grievances alone are insufficient to trigger protest; instead, individuals require access to mobilizing resources–such as communication channels and networks–to translate discontent into collective action.

This article demonstrates that all three resource dimensions–societal organizations, online networks, and financial means–consistently increase the protest incidence. However, while both societal organizations and online networks facilitate protest mobilization individually, their effects appear to be mutually substitutive. This finding supports the hypothesis formulated by Shirky (2011) that online connectivity diminishes the importance of traditional social movement organizations (SMOs). It also extends the claim by Earl and Kimport (2011) that SMOs become less important for online protest by showing that their relevance may also decline in offline contexts. Additionally, the results suggest that societal organizations play a key mediating role in translating local grievances into protest activity, contributing to the broader literature on the mediating effects of grievances through political mobilization (Kurer *et al.*, 2019), socio-economic environments (Dalton *et al.*, 2010), and macroeconomic and political conditions (Grasso and Giugni, 2016).

The first empirical challenge lies in the appropriate operationalization of the three resource dimensions. To address this, I employ a principal component analysis (PCA) on a set of variables commonly associated with each dimension. The PCA yields three distinct and interpretable components, corresponding to societal organizations, online networks, and financial means. I then estimate the baseline empirical model using a Poisson quasi-maximum likelihood estimator (QMLE), which provides estimates of the relative effect sizes of the three resource dimensions, as well as their interactive effects, particularly between societal organizations and online networks. Additionally, another empirical specification incorporates an indicator for grievances, constructed from measures of perceived discrimination, income dissatisfaction, and inequality. This indicator is used to examine the mediating role of resources in translating grievances into protest activity.

This article contributes to several strands of the literature on the determinants of protest activity. First, disentangling the effects of distinct resource dimensions advances our understanding of which types of resources are most relevant for protest coordination. Second, it engages with the growing body of research on online mobilization (Enikolopov *et al.*, 2020; Larson *et al.*, 2019; Mundt *et al.*, 2018) and broader network effects (Bursztyn *et al.*, 2021; Cantoni *et al.*, 2019). In contrast to these studies, this article focuses on how geographically defined network connectivity and depth affect protest outcomes. Third, the analysis of organizational structures contributes to the literature on the role of NGOs and civil society organizations in protest coordination (Murdie and Bhasin, 2011; Gillham *et al.*, 2019), with a particular emphasis on the substitutive relationship between societal organizations and online networks. Finally, the article connects to work by Kurer *et al.* (2019), Dalton *et al.* (2010), and Grasso and Giugni (2016), among others, by examining how grievance-based protest is mediated by contextual resource availability.

This article is organized as follows. Section 2.2 characterizes the three resource dimensions under investigation and derives five testable propositions. Section 2.3 describes the operationalization procedure using a principal component analysis and describes the empirical identification strategy. Section 3.4 presents the results and their interpretation, followed by some concluding remarks.

2.2 Resources and Protest Mobilization

Assessing the effects of societal organizations, online networks, and financial means on protest incidence constitutes the primary objective of this study. In doing so, this work contributes to a literature stream rooted in Resource Mobilization Theory (RMT), which emphasizes the strategic role of resources in enabling collective action. RMT seeks to explain the onset and success of social movements based on the endowments of tangible and intangible resources. Early RMT advocates like McCarthy and Zald (1977), Zald *et al.*

(1987), and Jenkins (1983) challenged the prior belief that social movements are a function of (relative) deprivation. Such views highlighted the role of grievances due to unmet individual expectations about economic or political affairs (Gurr, 1968, 1970) or one's disadvantaged relative position in society (Morrison, 1971). In contrast, RMT focuses on (rational) decision makers whose capabilities of organizing protests depend on access to and control over useful resources. Within this literature, the role of societal organizations, e.g., through the existence of social movement organizations (SMOs), became crucial as these institutions control large parts, but not all, of the available resources (Opp, 2009).

By shifting the focus from individual or structural grievances to resources like societal organizations or financial endowments, RMT potentially solves the free-rider problem more plausibly than grievance/deprivation-based theories. Consistent with Olson (1971, p. 51), free-riding can be avoided if large groups are capable of creating selective incentives at the individual level. Organizational resources like networks or communication channels may create such individual-level incentives for protest participation. Accordingly, Jenkins (1983, p. 538) argues that "the mobilization potential of a group is largely determined by the degree of preexisting group organization." The interplay between such societal organizations and grievances will be one of the formulated hypotheses below.

In RMT, resources are generally defined in a very broad sense, which becomes apparent in the definition of Gamson *et al.* (1982, p. 23): "By resources, we mean those objects which can be used by the group to achieve its collective goals." To conceptualize resources, Edwards and McCarthy (2004) categorize resources into five types: material, social-organizational, human, cultural, and moral. Examples of each type are financial endowments, communication channels, protest entrepreneurs, tacit knowledge, and a movement's legitimacy in society, respectively. Fuchs (2006) names further material resources, e.g., money, organizations, means of communication, control over media, and non-material resources, e.g., loyalty, networks, personal connections, and public attention.

The large variety of resource types allows for a wide operationalizability of RMT. Within this literature, this study seeks to advance the understanding of specific resource types in protest mobilization by differentiating three distinct categories: organizations, online networks, and financial means. Doing so enables an assessment of their relative effect sizes and offers insights into potential interactive effects between resource dimensions and grievances. This disaggregation refines the theoretical understanding through which mobilization processes are driven by different resource dimensions and advances the empirical evaluation of how distinct resource endowments mediate grievances into collective action.

2.2.1 Societal Organizations

Societal organizations play a vital role in sustained protest activity. They produce resources such as communication infrastructure, recruit volunteers, and coordinate protest events (Edwards *et al.*, 2018). Additional functions include the collaboration among SMOs and with other societal organizations, which has the potential to increase a movement's visibility and mobilization size (Edwards *et al.*, 2018; Wang and Soule, 2012). Furthermore, higher societal organizations are linked to better coalition forming across movements (Van Dyke and Amos, 2017) and the aggregation of other resources like media visibility and peer influence (Pilny *et al.*, 2014). Regarding individual participation, Somma (2010) finds that active monetary and time contributions to SMOs significantly increase an individual's willingness to join protest events. These findings align with civil society theory, highlighting societal groups' network dimension and their mobilization potential for political participation (Putnam, 1995; Verba *et al.*, 1995).

Generally, societal organizations, e.g., non-governmental organizations (NGOs), are identified in the empirical literature as structures conducive to protest mobilization. Gillham *et al.* (2019) find that preexisting community resources facilitated the formation of SMOs in the case of the 2011 Occupy movement in the US. Other studies rely on metrics like union and church membership share or the density of such organizations. Dalton *et al.* (2010) proxy the degree of societal organization with membership shares of unions and churches and find positive effects on protest mobilization in a cross-country sample. Trade unions are found to be especially effective in protest mobilization in bad macroeconomic contexts using individual-level data from Europe (Quaranta, 2018). In another study based on European survey data, Kurer *et al.* (2019) find that individual union membership increases one's likelihood of engaging in protesting. Aggregated evidence is presented by Kim and McCarthy (2016), who find that higher religious density leads to increased protests at the US county level for various protest movements from 1960 to 1995. Taken together, the empirical evidence lends strong support to theoretical arguments emphasizing the positive effect of societal organizations in protest mobilization processes.

Building on this literature, this study predicts that resources fostered by societal organizations facilitate protest mobilization. Accordingly, Hypothesis 1a posits that a higher density of societal organizations leads to increased protest activity.

Hypothesis 1a. *Within a US county, a higher density of societal organizations leads to an increased protest incidence.*

2.2.2 Online Networks

Over the past two decades, online networks have received substantial scholarly attention for their role in enabling protest mobilization. These technologies influence protest participation through two primary mechanisms: (1) by providing decentralized channels for information dissemination, and (2) by reducing the costs of coordination (Zeitzoff, 2017). The significance of these mechanisms became particularly evident during the Arab Spring, which underscored the mobilizing potential of information and communication technologies (Earl, 2018). The first mechanism via enhanced information dissemination is illustrated by Earl *et al.* (2013), who show that protesters actively used Twitter during the Pittsburgh G20 meetings to share real-time updates about protest locations and police presence. Similarly, González-Bailón *et al.* (2011) identify the critical role of early participants in the Spanish protests, who disseminated information through social networks, thereby activating broader participation. The second mechanism via lower coordination costs is supported by Kadivar *et al.* (2025), who find that online protest calls during the Arab Spring led to increased protest activity in Iran. Likewise, Steinert-Threlkeld *et al.* (2015) report a positive correlation between the use of movement-relevant Twitter hashtags and protest activity during the same period.

More recent studies seek to assess individual positions in networks or exploit exogenous variation to demonstrate the importance of online networks for protest participation. Relying on Twitter data, Larson *et al.* (2019) find that protest participation during the Charlie Hebdo protests in France depended mainly on an individual's exposure to the intentions of one's online network. Similar findings are presented by Anduiza *et al.* (2014) for protests in Spain, where personal contact and online social networks were found to be the main mobilization channels. Enikolopov *et al.* (2020) present causal evidence based on time-variation in the adoption of the VK-network in Russia to assess online network effects in protest participation. The authors attribute most of the online network effect to reduced coordination costs instead of enhanced information spreading.

This study investigates the general effect of online network penetration and network depth on protest activity. This inquiry serves two primary purposes. First, unlike societal organizations, online networks function as decentralized communication and coordination infrastructures, operating without formal hierarchies. As a result, mobilization patterns facilitated by online networks may follow dynamics distinct from those associated with more hierarchical organizational forms. Second, much of the existing empirical literature focuses on specific protest movements, time periods, or platforms. In contrast, this study examines broader patterns of connectivity and the structural strength of online clusters. It seeks to understand how varying degrees of online network integration influence protest mobilization.

Building on these considerations, Hypothesis 1b posits that online networks increase the protest incidence by reducing coordination costs and facilitating the dissemination of protest-relevant information.

Hypothesis 1b. *Within a US county, greater online connectivity and network depth lead to an increased protest incidence.*

2.2.3 Financial Endowments

Financial endowments affect the level of protest activity both indirectly by improving the financial situation of societal organizations and directly by altering individual incentives. SMOs and societal organizations generally have a deep-rooted interest in aggregating a variety of resources, including financial. Common strategies for acquiring such resources include co-optation and patronage (Edwards *et al.*, 2018). Financial endowments are especially effective due to their fungibility. As Edwards *et al.* (2018) argue, financial endowments facilitate protests by enabling organizations to obtain the complementary assets they most urgently require. They also offer straightforward solutions to logistical challenges faced by protest organizations, such as securing office space, maintaining communication infrastructure, or launching advertising campaigns. Moreover, financial support can reduce the personal risks of participation by covering legal expenses, fines, or bailouts in the event of arrest (Andrews and Biggs, 2006).

Financial endowments also influence protest activity at the individual level by shaping participation incentives. In a seminal work, Brady *et al.* (1995) show that better situated individuals are more likely to engage in protest and electoral participation in the US. While higher income or wealth may raise the opportunity cost of participation (e.g., through forgone wages or leisure), other factors, such as increased autonomy and political efficacy, appear to dominate, resulting in a net positive effect of affluence on participation. Recent evidence by Miller *et al.* (2023) further suggests that non-liquid assets, in particular, are associated with higher protest activity. One possible mechanism is that local asset ownership influences individuals' risk preferences and strengthens their sense of community involvement. This interpretation aligns with findings that homeowners are more likely to volunteer (Rotolo *et al.*, 2010) and to vote in local elections (McCabe, 2013), both based on US data.

The presented evidence points to a positive effect of income and wealth on the protest incidence in a given region. This relationship is documented in the literature at both the individual and aggregated levels. As a consequence, the resulting hypothesis theorizes that financial endowments should increase the protest incidence at the US county level.

Hypothesis 1c. *Within a US county, more financial endowments lead to an increased protest incidence.*

2.2.4 Intersection between Organizations and Online Networks

The rise of ICTs has significantly transformed patterns of protest mobilization, reshaping the relationship between SMOs and decentralized online networks (Earl, 2018). Traditionally, SMOs played a central role in reducing information costs and fostering network ties to facilitate collective action. ICTs, ranging from basic tools like email lists to more advanced platforms such as social media, replicate these functions while enabling more decentralized forms of coordination. Beyer (2014) argues that informal, loose networks are much more effective in organizing and coordinating protest than traditionally believed in resource-based protest theories. The expansion of online networks also blurs the distinction between organizers and participants, as mobilization increasingly occurs through decentralized, peer-driven channels (Earl, 2015). Bennett and Segerberg (2012) even argue that protest coordination via online channels may follow a fundamentally different logic, as it relies less on traditional organizational resources and more on personal content sharing in what the authors call "connective action."

These shifts in protest mobilization dynamics have led scholars to identify tensions between traditional SMO-based mobilization and emerging forms of network-driven, digitally mediated protest. In an early contribution, Shirky (2008) argues that ICTs not only diminish the centrality of SMOs in organizing protest but may also destabilize their role altogether. A shift of the importance in favor of online networks and communities, going together with decreased importance of SMOs, is also advocated by Earl (2015) for cases of online protesting. However, the authors acknowledge the remaining importance of SMOs for offline protest mobilization. Earl and Kimport (2011) also emphasize the declining relative importance of SMOs and call for further research into the conditions under which these organizations continue to play a meaningful role in contemporary protest dynamics.

While both societal organizations and online networks are theorized to independently enhance protest mobilization (as stated in Hypotheses 1a and 1b), empirical evidence on their joint effect remains limited. The reviewed literature suggests a potential crowding-out dynamic, whereby the presence of strong online network resources may diminish the marginal effect of societal organizations, and vice versa. This possibility motivates Hypothesis 2, which posits a negative interaction between societal organizations and online networks in shaping protest activity.

Hypothesis 2. *Within a US county, the impact of societal organizations and the depth of online networks on the protest incidence exhibits a crowding-out effect, whereby an* *increase in one reduces the effect size of the other.*

2.2.5 Intersection between Resources and Grievances

Grievance-based theories, widely discussed in the literature, attribute the emergence of protest to unfulfilled expectations regarding social, political, or economic conditions (Opp, 2009). When grievances are high within a social group, individuals are more likely to engage in protest, as the perceived opportunity costs of participation are lower and the expected benefits from potential policy change are greater. A substantial body of empirical research supports the role of grievances in driving protest activity. For instance, Ortiz *et al.* (2013) analyze protest events across more than 80 countries and identify four primary sources of grievance-based mobilization: failures of political representation, global injustice, violations of rights, and economic injustice, including austerity measures. Additional studies provide evidence for the role of economic grievances in particular. Using data from the World Values Survey, Jo and Choi (2019) find that unmet preferences for redistributive policies are associated with increased protest participation at the individual level. Similarly, Justino and Martorano (2019) document this relationship in a sample of Latin American countries. Dalton *et al.* (2010) report a modest positive association between individual-level grievances and protest behavior in a cross-national hierarchical model. In a related study using the European Social Survey, Kurer *et al.* (2019) find that increases in perceived deprivation are positively linked to protest participation. Economic grievances have also been shown to be a key driver of protest activity in the Middle East, particularly among students (Shafiq *et al.*, 2014).

Despite their prominence, grievance-based approaches to protest mobilization remain contested. A common critique is that these theories overemphasize individual-level discontent while neglecting the organizational and structural resources necessary for collective action. For example, Caren *et al.* (2011) find no significant relationship between personal income and protest participation in a large US sample, whereas factors such as union membership and residential location are significant predictors. Similarly, Welzel and Deutsch (2012), using data from the World Values Survey, show that personal values, rather than grievances, are more strongly associated with political participation. Historical evidence also underscores the importance of mobilizing structures: Andrews and Biggs (2006) identify organizational resources as the primary driver of protest activity during the 1960s sit-ins against racial discrimination. More recent work seeks to highlight the importance of political circumstances. For instance, Grasso and Giugni (2016) integrate grievance-based and context-based theories by demonstrating interactive effects between individual-level grievances and broader macroeconomic and political opportunity structures.

Given the contested empirical support for grievance-based theories, I argue that the effect of grievances on protest mobilization is critically mediated by the availability of mobilizing resources. The underlying rationale is that, in the absence of sufficient resources, such as organizational infrastructure or digital communication networks, grievances alone may not translate into collective action, as individuals lack the means to coordinate to express dissent. However, when resource endowments are sufficiently high, grievances can serve as a powerful catalyst for mobilization, as dissatisfaction with the status quo is more effectively channeled into protest activity. This logic motivates Hypothesis 3, which states that grievances increase protest activity only in contexts where resource availability is high. Finding evidence that the effect of grievances is mediated by resource availability contributes to reconciling the mixed empirical findings in the existing literature. Such a result provides an explanation of why grievances are associated with increased protest activity in some contexts but not in others. Specifically, it suggests that grievances alone are insufficient to drive mobilization unless individuals or groups also possess the necessary resources, organizational, digital, or other, to act upon their discontent.

Hypothesis 3. *Within a US county, the effect of grievances on the protest incidence is mediated by the level of available resources.*

2.3 Empirical Strategy

The empirical strategy proceeds in two steps. The first challenge is to appropriately operationalize the three resource dimensions: societal organizations, online networks, and financial endowments. This is addressed by applying PCA to construct composite indices for each dimension. Based on these indices, the following two models are developed to test Hypotheses 1a through 1c. The first specification focuses on the hypothesized substitutive relationship between societal organizations and online networks (Hypothesis 2), while the second examines the mediating role of resources in the relationship between grievances and protest activity (Hypothesis 3).

2.3.1 Measuring Resource Dimensions

In order to operationalize the three resource dimensions, societal organizations, online networks, and financial means, nine variables are selected to capture distinct aspects of each dimension (see Table 2.1). These variables are analyzed using a PCA. PCA is particularly well-suited for this task because it allows for categorizing variables into composite indicators that summarize the shared variance among correlated variables, thereby enhancing interpretability and reducing multicollinearity in subsequent regression analysis (James *et al.*, 2013). The resulting categorization is used to construct three composite indicators – one for each resource dimension. The empirical strategy then proceeds by specifying the regression model used to assess the relationship between these resource indicators and protest activity.

Societal organizations are measured through the density of civic organizations, which constitute one mechanism for enabling collective action. The primary indicator is the number of NGOs per capita, a widely used proxy for local organizational infrastructure in studies of protest and civic engagement (Murdie and Bhasin, 2011; Boulding, 2010). To broaden this measure, two additional variables are included: the number of charitable organizations per capita and a normalized count of civic organization pages on Facebook.[2] These variables capture both formal and digitally visible civic organizations that enhance mobilization potential by providing resources like communication channels, networks, and coordination infrastructure.

In the context of online networks, existing literature has largely concentrated on specific protest movements, digital platforms, or time periods. This study contributes to that body of work by introducing a regionally defined measure of network penetration and examining its association with protest activity at the county level. The operationalization includes three variables: (i) the relative likelihood that a Facebook user has a friendship link with another user in the same county, compared to the general probability of such links (network connectivity); (ii) the presence of network clusters, capturing the tendency for mutual connections among shared contacts; and (iii) the support ratio, which reflects the density and cohesion of local social ties. Together, these measures aim to capture both the online network connectivity and depth within local communities.

Financial endowments are captured using three variables that reflect different aspects of local economic resources. The first is the logarithm of the median income, which serves as a measure of the typical resident's financial standing. The second is the share of individuals with non-wage income, used as a proxy for accumulated wealth and financial diversification. Finally, the logarithm of income per capita captures the overall level of economic resources available within a county. Together, these indicators provide a comprehensive view of the monetary endowments that may influence the capacity and willingness of individuals to participate in protest activity.

PCA is an exploratory factor analysis technique that reduces the dimensionality of a

[2]The thrid variable, normalized count of civic organization pages on Facebook, is especially useful in disentangling the effect of societal organizations and online networks as it proxies the degree to which societal organizations use online communication channels. In contrast, online networks differ as they measure the general degree of connectivity and thereby provide a decentralized mechanism for protest coordination.

Table 2.1: Description Resource Variables

Variable	Description & Source
Organizational Capacities	
NGOs p.c.	Number of Non-Governmental Organizations per capita. Source: NCCS (2024)
Charities p.c.	Number of charitable organizations per capita. Source: NCCS (2024)
Civic organizations	The estimated count of Facebook Pages classified as "Public Good" based on page title, category, and other characteristics per 1,000 users within a county. Source: Chetty *et al.* (2022)
Online Networks	
Network connectivity	The likelihood that a given Facebook user has a friendship link with another user residing in the same county, relative to overall friendship probabilities. Source:
Network clusters	The proportion of an individual's friendship pairs in which both friends are also connected to each other. Source: Chetty *et al.* (2022)
Support ratio	The share of within-county friendships in which both individuals have a mutual third friend residing in the same county. Source: Chetty *et al.* (2022)
Financial Resources	
Log median income	Logarithm of median income per household. Source: U.S. Census Bureau (2022)
Wealth	Share of individuals earning non-wage income. Source: U.S. Census Bureau (2022)
Log income p.c.	Logarithm of county income per capita. Source: U.S. Census Bureau (2022)

Note: All variables are collected on the US county level. The data retrieved from Chetty *et al.* (2022) are measures as of May, 2022. The Bailey *et al.* (2018) data are available for 2018 only. However, Bailey *et al.* (2021) demonstrate their persistency and usefulness over time. All remaining variables are collected annually.

large set of variables while preserving as much of the original variance as possible (James *et al.*, 2013). This is particularly valuable in the social sciences, where complex constructs often cannot be adequately captured by a single observable variable. PCA enables researchers to derive composite indicators that reflect underlying latent dimensions, based on the covariance structure of the data, rather than subjective assumptions (Kabacoff, 2022; Jolliffe and Cadima, 2016). A key advantage of PCA is that it avoids multicollinearity and identification issues by allowing the inclusion of multiple, potentially correlated variables without requiring a priori assignment to specific sub-concepts. This reduces the risk of omitted variable bias and enhances the objectivity of the operationalization process. Recent applications underscore its utility, for instance, Vâlsan *et al.* (2023) apply PCA to measure social capital across US counties in order to reduce reductionism and arbitrariness in measuring complex social phenomena.

In a PCA, principal components (PCs) are computed based on the covariance structure of the underlying variables (James *et al.*, 2013). PCs are linear combinations of all considered variables k. The weights of the linear combination for the first PC are chosen so that the first PC's variance is maximized. The same procedure is applied to all subsequent PCs, plus a restriction that all subsequent PCs must be orthogonal to the previous PCs. Finally, the number of k principal components (PCs) to include in the subsequent analysis can be determined by the researcher. For better interpretability, the PCA is conducted using the oblique (promax) rotation.

Interpreting the PCA results in Table 2.2, the variables cluster according to the previous classification of the three resource dimensions, demonstrating the appropriateness for building indicators for these three resource dimensions. This clustering is indicated by the factor loadings of all variables on the first three PCs. For the hypothesized resource categories, the loadings of each variable are at least 0.5 for their respective resource dimension, indicating strong weights of these variables for the PC. The three PCs explain jointly 76% of the overall variance of all variables, and their eigenvalues exceed 1. These results demonstrate the homogeneity of the formulated resource dimensions and build the data-driven operationalization for the subsequent analysis of each dimension.

Table 2.2: Principal Component Analysis Results

Principal Components	PC1	PC2	PC3
Resource Dimensions	Online Networks	Financial Res.	Organizations
Organizations			
NGOs p.c.	0.38	0.32	**0.50**
Charities p.c.	-0.19	-0.13	**0.91**
Civic organizations	-0.05	0.02	**0.82**
Online Networks			
Network connectivity	**0.79**	-0.13	0.24
Network clusters	**0.89**	-0.02	-0.16
Support ratio	**0.84**	0.00	-0.19
Financial Resources			
Log median income	-0.38	**0.77**	-0.13
Wealth	0.20	**0.89**	0.09
Log income p.c.	-0.35	**0.79**	0.01
SS loadings	2.69	2.25	1.87
Proportion Var	0.30	0.25	0.21
Cumulative Var	0.30	0.55	0.76

Note: The displayed numbers represent factor loadings of the PCA using oblique (promax) rotation for the year 2022. All loadings larger than absolute value of 0.5 are highlighted in bold. All variables enter the PCA in a standardized form.

For ease of interpretation, the final indicators are constructed as arithmetic means of the standardized variables associated with each principal component. This step is deemed helpful as some variables exhibit non-negligible loadings on multiple components, making direct use of the PCA scores less transparent. A similar approach is adopted by Durante *et al.* (2025) in their construction of social capital indicators. Notably, the resulting composite indicators are highly correlated with the corresponding principal components (correlation coefficients exceed 0.97), indicating that the simplified arithmetic means retain nearly all the explanatory power of the PCA-derived scores. Consequently, the subsequent analysis relies on these averaged indicators to enhance interpretability.

2.3.2 Identification Strategy

Having obtained three indicators for societal organizations, online networks, and financial means, I proceed in the empirical strategy by describing the regression model for assessing their impact on protest activity. A first set of regressions assesses the general effects of each resource dimension on the protest incidence (Hypotheses 1a-1c) and the interaction between societal organizations and online networks (Hypothesis 2). Then, a second set of regressions investigates the mediation of grievances via available resources (Hypothesis 3). The presented evidence is complemented by a sensitivity analysis for different protest motivations and alternative estimation techniques.

The analysis focuses on the incidence of protest events at the US county level. Protest data are sourced from the *Crowd Counting Consortium* (CCC), a widely used event database in academic research (Dorff *et al.*, 2023), and aggregated to the county level. The CCC compiles protest events through automated web crawls, systematic reviews of social media and organizational websites, Google searches, and public submissions. Between 2018 and 2022, the database records 125,344 protest events in the US. While the CCC also provides estimates of participant counts, these figures are inconsistent as lower and upper bounds vary substantially and are missing for more than half of the events. Accordingly, the analysis focuses on the event incidence rather than the crowd size. As a robustness check, the empirical models are re-estimated using data from the Armed Conflict Location and Event Data Project (ACLED) (Raleigh *et al.*, 2010).

To assess the effect of the resource dimensions on the incidence of protests, I estimate the following Poisson count model:

$$ProtestEvents_i = \beta_0 + \beta_1 Org_i + \beta_2 Net_i + \beta_3 Mon_i + \beta_4 (Org_i \times Net_i) + \beta_{5-19} X_i + \epsilon_i \quad (2.1)$$

The primary explanatory variables are Org_i, Net_i, and Mon_i, which measure the resource dimensions of societal organizations, online networks, and monetary resources, respectively. All variables are standardized to facilitate interpretation of coefficients as standardized (beta) effects (Hypotheses 1a–1c). The interaction term $Org_i \times Net_i$ tests Hypothesis 2, which predicts a negative coefficient, reflecting the expectation that the marginal effect of one resource diminishes as the other increases.

The matrix X_i includes a set of control variables to account for potential confounders. Grievance-related factors include the number of fatal police encounters involving Black individuals (Mapping Police Violence, 2025), the unemployment rate, and the Gini coefficient. Socio-demographic controls include log population, population density, the share of individuals without a school degree, the number of college students per capita, the share of post-1990 immigrants, the population share of people aged between 20 and 39, and the share of White and Black individuals. Urbanicity is captured by two dummy variables for high-density areas (density > 40,000 and > 20,000 per square mile). To account for political opportunity structures, the closeness of the most recent presidential election (losing party votes divided by winning party votes) is included, based on data from MIT Election Data and Science Lab (2018). Unless otherwise noted, all variables are drawn from the U.S. Census Bureau (2022).

Equation 2.1 is estimated using a Poisson quasi-maximum likelihood estimator (QMLE) for different specifications. This estimation technique is suitable for count data, which includes a number of zeros and is skewed towards high values. Even if the dependent variable is not Poisson distributed, the estimation will be consistent and maintains some efficiency properties (Wooldridge, 2010). The baseline year for the estimation is 2022. To demonstrate inter-temporal robustness, other specifications for the years 2018 and 2020 are included, too. Another specification trims the sample to only include counties with a population size of at least 30,000 to exclude small US counties for 2022. The last specification replaces the dependent variable from the CCC database with the protest incidence in the ACLED database to test robustness to a different data source. Further robustness to methodological choices is established by employing alternative estimation techniques.

The second regression set is constructed to test the mediated effect of grievances via resources (Hypothesis 3). To this end, the model in Equation 2.2 includes an interaction term between a grievances measure and resources. The model is, again, estimated with Poisson QMLE, where the same set of controls is employed as in the previous set of regressions.

$$ProtestEvents_i = \beta_0 + \beta_1 Org_i + \beta_2 Net_i + \beta_3 Mon_i + \beta_4 Gri_i + \beta_5 (Res_i \times Gri_i) + \beta_{6-16} X_i + \epsilon_i \tag{2.2}$$

The Gri_i variable is the arithmetic mean of the three standardized grievance variables, which include the number of fatal police encounters with a black individual, the unemployment rate, and the Gini coefficient. These grievance dimensions are chosen as they cover the most salient protest topics during the time frame under investigation: race-based discrimination, economic dissatisfaction, and relative deprivation. The model is then estimated, replacing Res_i with each of the resource dimensions separately, then with three interaction terms at once, and finally with an overall resource indicator which is formed by the arithmetic mean of the individual resource dimensions. Thus, it becomes feasible to assess the interactive effects of each resource dimension separately, at once, and jointly as one indicator.

2.4 Results

The results are presented in three steps, progressing from the effects of resources to the mediation of grievances. The first subsection examines the relationship between the three resource dimensions and the incidence of protest events. The second subsection explores the theorized mediating role of resources in the effect of grievances on protests. The final subsection investigates heterogeneity in effects across motivation-specific protests and includes a series of robustness checks.

2.4.1 Effects of Resources on Protest Events

The results of the first set of estimations highlight the significance of all three resource dimensions in protest coordination. Table 2.3 reports a subset of the estimation results, excluding control variables for brevity. The full regression output, including all controls, is provided in Appendix Table 2.6. Importantly, the regression results are interpretable as incident rate ratios (IRR), implying that values below 1 indicate a negative effect of the variable on the protest incidence in a given county. The IRR is the exponentiated Poisson QMLE point estimate, $exp(\beta_{Poisson})$, and is interpretable as a percentage change in the dependent variable due to an increase in the independent variable.

The presented evidence is in line with Hypothesis 1a-1c as each resource dimension has a positive impact on the protest incidence. For instance, an increase of societal organizations by one standard deviation increases the number of expected protest events by 24.6% in 2022. This effect is robust to excluding all counties with a population smaller than 30,000, the alternative dependent variable from ACLED, and the years 2020 and 2018. The effects of online networks and monetary means are also statistically identified for all specifications with similar effect sizes. These findings underscore the significance of all three resource dimensions in shaping protest incidence and offer strong support for Hypotheses 1a-1c.

The hypothesized substitutive effect between societal organizations and online networks is also supported by the empirical results (Hypothesis 2). The interaction term is smaller than 1, indicating that for high values of societal organizations, the effect of online networks decreases and vice versa. This effect is identified for all specifications except the one with the trimmed sample. To quantify this effect, I calculate the conditional marginal effect that is defined as the derivation of the conditional mean function

Table 2.3: IRR effect of different resource dimensions on protest events.

	(1) 2022	(2) Pop Cap	(3) ACLED	(4) 2020	(5) 2018
R: Organizations	1.246***	1.580**	1.248***	1.245***	1.243***
	(0.072)	(0.296)	(0.065)	(0.063)	(0.061)
R: Online Network	1.220**	1.230**	1.219***	1.232***	1.089*
	(0.121)	(0.125)	(0.079)	(0.075)	(0.050)
R: Monetary Means	1.439***	1.471***	1.133*	1.121*	1.111**
	(0.153)	(0.168)	(0.075)	(0.067)	(0.056)
R: Org. $\times$ R: Network	0.925**	1.034	0.899***	0.926***	0.934***
	(0.033)	(0.125)	(0.024)	(0.026)	(0.024)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	3036	1405	3036	3043	3043
Pseudo-R^2	0.754	0.695	0.842	0.815	0.831

Heteroskedasticity robust SE in parentheses; * p <0.1, ** p <0.05, *** p <0.01

Note: The exponentiated Poisson QMLE coefficients represent IRR for the model presented in Equation 2.1. Model (2) estimates the effects on protest events in a trimmed sample for counties with a population of > 30,000 in 2022. Therefore, the number of observations is smaller. Specification (3) replaces the database for the protest count variable with ACLED for 2022. Specifications (4) and (5) present the estimates of the model presented in Equation 2.1 for the years 2020 and 2018, respectively.

with respect to the variable of interest and plug in different values for X.[3] Based on Model (1), a one-unit increase in organizational resources is associated with an increase of 0.48 in the expected number of protests when online network presence is low (at the 5th percentile), compared to an increase of 0.29 when online network presence is high (at the 95th percentile). A similar pattern holds for the effect of online networks: at low levels of societal organizations (5th percentile), a one-unit increase in online network strength increases protest activity by 0.43, whereas the effect diminishes to 0.15 at high levels of societal organizations (95th percentile). These findings support Hypothesis 2, as a larger presence of central, hierarchical organizations reduces the effect of decentralized, non-hierarchical online networks and vice versa.

The findings underscore the enduring relevance of resources in protest mobilization. Among these, societal organizations have traditionally played a central role and have been extensively examined in the literature (Andrews and Biggs, 2006; Edwards *et al.*, 2018). Building on a growing body of research on digital coordination (Larson *et al.*, 2019; Kadivar *et al.*, 2025), this article highlights the rising importance of online networks in facilitating collective action. A novel contribution of this study is the identification of a substitutive relationship between societal organizations and online networks. Whereas prior empirical work has typically modeled resource dimensions additively (e.g., Gill-

[3] The derivative of the conditional mean function with respect to the variable of interest x_j is $\frac{\partial E(y|\mathbf{x})}{\partial x_j} = exp(\mathbf{x}\beta)(x_i\beta_{i,j} + \beta_j)$, where the interaction term is $x_i \times x_j$ and its point estimate is $\beta_{i,j}$.

ham *et al.*, 2019; Castelli Gattinara *et al.*, 2022), the results presented here suggest that more flexible specifications may better capture the interactive dynamics of resource-based protest mobilization.

The findings have practical implications for patterns of protest and political participation, particularly in economically disadvantaged regions. Such regions often exhibit weaker socio-economic institutions and lower levels of societal organizations – what civil society theory refers to as diminished social capital (Putnam, 1995; Verba *et al.*, 1995). The empirical results suggest that in the absence of a strong organizational infrastructure, higher levels of online connectivity can serve as a substitute. This dynamic may be especially relevant in structurally disadvantaged areas, where digital networks can potentially emerge relatively quickly and help mitigate participation gaps relative to more institutionally robust regions. Such dynamics may be especially relevant in democratizing countries where civil organizations are potentially weaker.

2.4.2 Mediation of Grievances

The next set of estimates tests the hypothesized mediation of grievances via resources. As presented in Section 2.2, there is mixed evidence on the direct effect of grievances on the incidence of protest events and individual participation in protest events. Some studies find a mediated effect via macroeconomic circumstances or the political context. In this study, the interaction between different dimensions of resources with the grievances level, measured as an indicator of police violence against black individuals, the Gini coefficient, and the unemployment rate, is of particular interest (Hypothesis 3).

The results reported in Table 2.4 indicate robustness of the resource effects and a mediation between resources and grievances. As visible in the first three lines, the resource indicators significantly affect the level of protests, also when including the interaction effect with grievances. These findings increase the credibility of the results obtained in the previous subsection. Also, the summarized indicator (RMT) has a significant effect on the incidence of protest events. This finding should be unsurprising since the indicator is composed of three highly significant sub-indicators for each resource dimension.

The mediation of grievances via resources is found on a general level and is likely to be driven by the dimension of societal organizations. The interaction term between societal organizations and grievances is highly significant and larger than 1, indicating that both variables reinforce each other. There is no evidence that online networks have such mediating effects on grievances, and only weak evidence for such an effect of monetary resources. Summarizing all resource dimensions to the single indicator RMT (using the mean of the three standardized variables) and interacting it with grievances leads to a significant effect of both variables separately and jointly in the interaction term.

Table 2.4: IRR effects of interactions between resource dimensions and grievances on protest events.

	(1) Baseline	(2) Org.	(3) Network	(4) Monetary	(5) All	(6) RMT
R: Organizations	1.308***	1.221***	1.308***	1.301***	1.224***	
	(0.068)	(0.067)	(0.068)	(0.067)	(0.068)	
R: Online Network	1.252**	1.361***	1.242*	1.266**	1.339**	
	(0.130)	(0.150)	(0.147)	(0.131)	(0.170)	
R: Monetary Means	1.497***	1.369***	1.500***	1.361***	1.360***	
	(0.139)	(0.129)	(0.135)	(0.140)	(0.129)	
Grievances	1.033	1.157***	1.053	0.975	1.179*	1.264***
	(0.049)	(0.033)	(0.117)	(0.051)	(0.114)	(0.082)
Gri. × R: Org.		1.205***			1.198***	
		(0.045)			(0.058)	
Gri. × R: Network			1.007		1.012	
			(0.046)		(0.032)	
Gri. × R: Monetary				1.089*	1.011	
				(0.052)	(0.039)	
RMT						1.502***
						(0.107)
Gri. × RMT						1.143***
						(0.042)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3067	3067	3067	3067	3067	3067
Pseudo-R^2	0.750	0.757	0.750	0.752	0.757	0.754

Heteroskedasticity robust SE in parentheses; * p <0.1, ** p <0.05, *** p <0.01

Note: The exponentiated Poisson QMLE coefficients represent IRR for the model presented in Equation 2.2 for the year 2022. The RMT variable is the arithmetic mean of the three resource dimensions.

The interaction between the resources indicator and grievances is analyzed more closely in Figure 2.1, highlighting their interdependencies. The figure displays the conditional marginal effects of grievances and resource variables, computed while holding all other covariates at their sample means and varying the variable of interest. The effect of grievances on the protest incidence is small and statistically insignificant for low values of resources: when resources are low (corresponding to their 5% percentile), a one standard deviation increase in grievances increases the number of expected protests by 0.07 with a standard error of 0.04. This effect increases substantially in size, implying that a one standard deviation increase in grievances for a higher value of resources (corresponding to their 95% percentile) increases the expected number of protests by 2.18. Similarly, the effect of resources also increases substantially in size for higher levels of grievances.

The evidence, thereby, supports Hypothesis 3 as the effects of grievances and resources reinforce each other substantially. This relationship is probably driven by the variable of societal organizations, implying that centrally organized institutions are the central

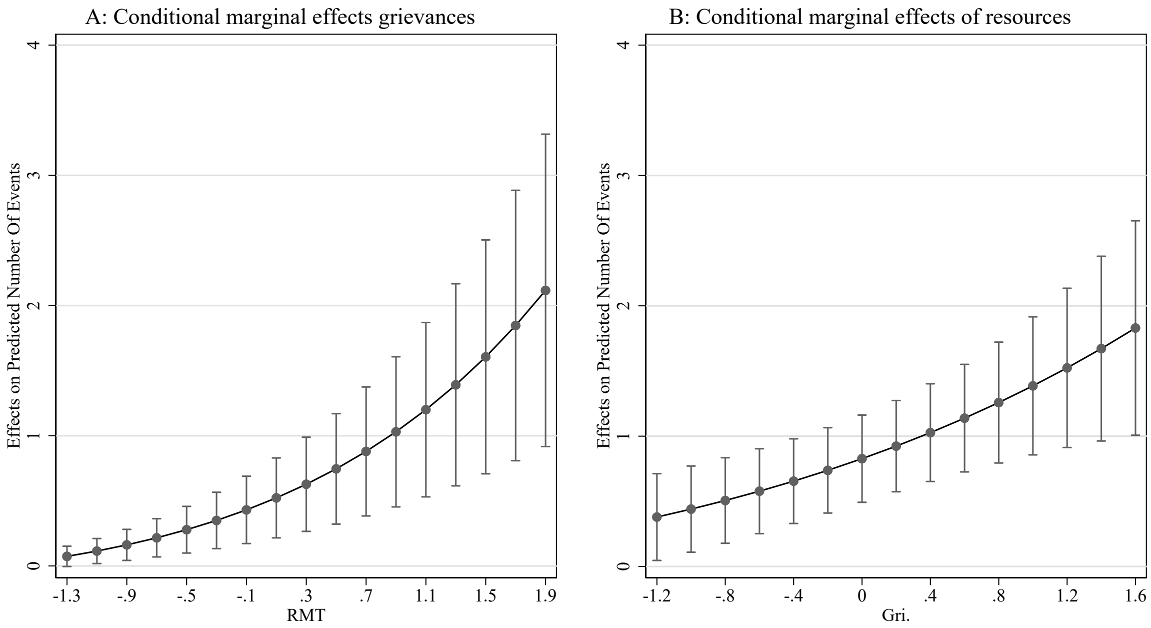


Figure 2.1: Marginal effects of grievances and resources on the protest incidence obtained from Model (1) in Table 2.4. The range of each x-axis spans from the 10th to the 90th percentile of the respective variable.

vehicle through which individuals are mobilized via their grievance levels. There is mixed evidence on the independent effect of grievances on the protest incidence. For those specifications with statistically significant interaction terms, Model (2), (5), and (6), the independent effect of grievances is identified (at least at the 10% level). This finding provides an explanation of why some studies report null effects of grievances on protest, while others find strong and statistically significant relationships. Consistent with the empirical results, the effect of grievances appears to be critically mediated by regional resource endowments. When models fail to adequately control for these resources or omit interactive terms, the independent effect of grievances may not be statistically identifiable.

The results, in combination with the existing literature, suggest that grievances alone are insufficient to trigger (substantial) protest activity. Rather, it is the presence of mobilizing resources – particularly societal organizations – that enables grievances to translate into collective action. From a collective action perspective, this may reflect the absence of selective incentives when grievances rise in isolation. Only when sufficient mobilizing resources are present do increased levels of grievances lead to a substantial increase in protest activity. This finding raises important questions for future research. In contexts where societal organizations are weak, do heightened grievances translate into alternative forms of political expression, such as protest voting, online radicalization, or non-political responses like increased criminal activity? While these questions lie beyond the scope of this article, they point to promising avenues for further investigation into the political consequences of unchanneled discontent.

2.4.3 Movement Heterogeneity and Robustness

The final set of estimation results is supposed to test the robustness of the previously presented findings. First, I disaggregate the dependent variable by specific protest issues to evaluate whether the estimated effects are identifiable across distinct political clusters and movements. This allows us to test the generalizability of the results beyond aggregate measures. The second part of the sensitivity analysis is a set of specifications that vary the estimation techniques and model setup. Generally, the findings are robust to a wide range of alternative specifications.

To assess whether the observed effects generalize across different types of protest events, the dependent variable is disaggregated into three issue-specific categories. These categories correspond to distinct reasons for protesting: discrimination-based grievances, public goods, and individual rights. The first category captures protests related to racism and policing – issues that gained high salience following the George Floyd incident in 2020. The second category includes protests concerning the provision of public goods, specifically environmental and educational policies. The third category focuses on individual rights, including protests related to gun rights and civil liberties, which typically aim to preserve or expand personal freedoms.

While the results generally demonstrate robustness for the effect of the three resource dimensions and their interaction, the online network effect seems to be sensitive to different protest issues at first. However, the significant negative interaction term between societal organizations and online networks indicates that online networks may only have a mobilizing effect when societal organizations are sufficiently small. Otherwise, their effect is crowded out by societal organizations. In the case of environmental protests, the online network effect is statistically significant and positive for low values of societal organizations (corresponding to the 5% percentile), while the variable's effect holding the interacted variable at the mean is statistically insignificant. In conclusion, disaggregating the effects of the three resource dimensions across protest-specific issues demonstrates the robustness of the main findings.

To address potential concerns regarding the econometric strategy, Table 2.8 in the Appendix presents a series of alternative model specifications and estimators. These include an OLS estimation of the baseline model (Equation 2.1) using an inverse hyperbolic sine (IHS) transformation of the dependent variable[4] as well as a negative binomial regres-

[4]The IHS transformation, defined as: $arsinh\ x = ln(x + \sqrt{x^2+1})$, is similar to the logarithmic

Table 2.5: Robustness of IRR effects for specific protest issues.

	(1) Racism	(2) Policing	(3) Environment	(4) Education	(5) Guns	(6) Civil Rights
R: Organizations	1.293***	1.321***	1.209**	1.255***	1.220***	1.319***
	(0.091)	(0.106)	(0.099)	(0.073)	(0.065)	(0.082)
R: Online Network	1.117	0.920	1.208	1.140	1.214**	1.255
	(0.214)	(0.196)	(0.183)	(0.108)	(0.094)	(0.181)
R: Monetary Means	2.073**	2.967***	1.819**	1.600***	1.552***	1.490*
	(0.643)	(0.791)	(0.496)	(0.231)	(0.247)	(0.338)
R: Org. × R: Net.	0.917**	0.943	0.880***	0.918**	0.938*	0.916***
	(0.039)	(0.045)	(0.038)	(0.032)	(0.032)	(0.030)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3036	3036	3036	3036	3036	3036
Pseudo-R^2	0.618	0.582	0.663	0.751	0.606	0.517

Heteroskedasticity robust SE in parentheses; * p <0.1, ** p <0.05, *** p <0.01

Note: The exponentiated Poisson QMLE coefficients represent IRR for the model presented in Equation 2.1, where the dependent variable is the protest count of specific protest issues in 2022.

sion. Across both specifications, the three resource dimensions generally show significant effects on protest activity. An exception arises in the negative binomial model, where the online network variable loses significance, potentially due to the moderating influence of the interaction term, as discussed previously.

Specifications (3) to (5) distinguish between the different margins of the resource effects on protest activity. Specifications (3) and (4) identify the effect of resources on the intensive margin. For this, only observations that showed at least one protest event in 2022 were included, and the model in Equation 2.1 is estimated using OLS (for an inverse hyperbolic transformed dependent variable) and using Poisson QMLE. Both specifications speak in favor of the identified effects of the three resource dimensions on protest activity, although the interaction term is barely identified. Specification (5) gives an indication of the effects on the extensive margin for which the dependent variable was transformed to be 1 if there was at least one protest event in 2022 and 0 otherwise. This specification highlights the importance of societal organizations and online networks in the process of coordinating the "first" protest event. Financial means, however, do not have a statistically identified impact, providing preliminary evidence that financial means may primarily affect protests on the intensive margin.

In conclusion, the sensitivity analysis supports the findings of the previous subsections. The theorized effects are consistently identifiable across a range of empirical specifications and hold for both the intensive and extensive margins of protest activity.

transformation but accommodates zero values

2.5 Conclusion

Motivated by the growing importance of online networks in contemporary collective action, this article examines the effects of three distinct resource dimensions–societal organizations, online networks, and financial endowments–on protest incidence. The analysis contributes to the literature on Resource Mobilization Theory by empirically disentangling the relative influence of these resource types. Each dimension is operationalized using principal component analysis (PCA), and the resulting indices are incorporated into a Poisson count model. The results indicate that all three resource dimensions are significant predictors of protest activity at the US county level.

A key contribution of this article is the argument that centrally organized, hierarchical organizations and decentralized online networks function as substitutes in mobilization processes. This claim is strongly supported by the empirical findings. The results carry important policy implications: as collective action becomes less reliant on centralized, hierarchical structures and increasingly driven by decentralized online communities, smaller and less formally organized groups may gain greater influence through their ability to coordinate via digital platforms. This shift underscores the growing role of information technology in collective action. Future research is needed to deepen our understanding of which groups gain influence through these changing contexts and which potentially lose due to shifting mobilization patterns.

Furthermore, this article argues that grievances are crucially mediated by resources. This argument is strongly supported by the empirical findings, particularly through the role of societal organizations, which are found to mediate the effect of grievances on protest activity. These results offer a potential explanation for the mixed empirical evidence regarding the impact of grievances on protest outcomes: studies that do not adequately account for this interactive effect may arrive at distorted conclusions. Future research could further disaggregate protest issues to examine whether the grievance effect systematically varies across different domains. Consistent with the findings presented here, resource endowments emerge as robust predictors of protest activity, supporting the view that they constitute structural factors facilitating collective action.

2.6 Appendix

Table 2.6: IRR effect of different resource dimensions on protest events.

	(1) Baseline	(2) Pop Cap	(3) ACLED	(4) 2020	(5) 2018
R: Organizations	1.246***	1.580**	1.248***	1.245***	1.243***
	(0.072)	(0.296)	(0.065)	(0.063)	(0.061)
R: Online Network	1.220**	1.230**	1.219***	1.232***	1.089*
	(0.121)	(0.125)	(0.079)	(0.075)	(0.050)
R: Monetary Means	1.439***	1.471***	1.133*	1.121*	1.111**
	(0.153)	(0.168)	(0.075)	(0.067)	(0.056)
R: Org. $\times$ R: Network	0.925**	1.034	0.899***	0.926***	0.934***
	(0.033)	(0.125)	(0.024)	(0.026)	(0.024)
Police Shooting B.	1.009	1.011	0.938***	0.988	0.982
	(0.049)	(0.052)	(0.021)	(0.027)	(0.020)
Unemployment	1.105**	1.117**	1.062**	1.079***	1.090***
	(0.046)	(0.057)	(0.028)	(0.031)	(0.022)
Gini	1.023	1.025	1.084***	1.049***	1.062***
	(0.021)	(0.026)	(0.015)	(0.010)	(0.008)
Log Population	3.127***	2.932***	3.404***	3.031***	2.912***
	(0.229)	(0.239)	(0.176)	(0.156)	(0.108)
Density	1.042	1.039	1.044	1.079**	1.022
	(0.053)	(0.055)	(0.035)	(0.033)	(0.027)
No school degree	0.226	0.631	0.015***	0.014***	0.002***
	(0.576)	(1.766)	(0.017)	(0.018)	(0.002)
College Students p.c.	4.544	2.413	2.703	2.212	2.642
	(7.421)	(4.182)	(2.701)	(2.765)	(2.767)
Share immigrants	0.058**	0.048**	0.377	0.278	1.228
	(0.081)	(0.072)	(0.279)	(0.223)	(0.678)
Share Aged 20-39	1.086***	1.087***	1.081***	1.051***	1.044***
	(0.016)	(0.016)	(0.011)	(0.012)	(0.011)
Pop. share white	0.786	0.811	0.420**	0.586	0.681
	(0.469)	(0.514)	(0.142)	(0.283)	(0.288)
Pop. share black	0.275*	0.312	0.379**	0.103***	0.131***
	(0.205)	(0.247)	(0.146)	(0.046)	(0.053)
Center 1	0.299***	0.372***	0.427**	0.371***	1.818***
	(0.071)	(0.132)	(0.148)	(0.067)	(0.358)
Center 2	3.872***	3.360***	1.470	2.837***	0.472***
	(0.931)	(1.025)	(0.442)	(0.558)	(0.093)
Election Closeness	0.739	0.698	0.891	0.780	1.029
	(0.170)	(0.164)	(0.134)	(0.152)	(0.133)
Observations	3036	1405	3036	3043	3043
Pseudo-R^2	0.754	0.695	0.842	0.815	0.831

Heteroskedasticity robust SE in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The exponentiated Poisson QMLE coefficients represent IRR for the model presented in Equation 2.1. Model (2) estimates the effects on protest events in a trimmed sample for counties with a population of > 30,000 in 2022. Therefore, the number of observations is smaller. Specification (3) replaces the database for the protest count variable with ACLED for 2022.

Table 2.7: IRR effects of interactions between resource dimensions and grievances on protest events.

	(1) Baseline	(2) Org	(3) Network	(4) Mon	(5) All	(6) RMP
R: Organizations	1.308***	1.221***	1.308***	1.301***	1.224***	
	(0.068)	(0.067)	(0.068)	(0.067)	(0.068)	
R: Online Network	1.252**	1.361***	1.242*	1.266**	1.339**	
	(0.130)	(0.150)	(0.147)	(0.131)	(0.170)	
R: Monetary Means	1.497***	1.369***	1.500***	1.361***	1.360***	
	(0.139)	(0.129)	(0.135)	(0.140)	(0.129)	
Grievances	1.033	1.157***	1.053	0.975	1.179*	1.264***
	(0.049)	(0.033)	(0.117)	(0.051)	(0.114)	(0.082)
Log Population	3.090***	3.227***	3.082***	3.104***	3.211***	3.143***
	(0.223)	(0.218)	(0.221)	(0.205)	(0.229)	(0.206)
Density	1.070	1.066	1.068	1.076	1.064	1.061
	(0.056)	(0.054)	(0.058)	(0.057)	(0.055)	(0.055)
No school degree	2.043	1.189	2.222	1.186	1.283	0.830
	(4.256)	(2.330)	(4.620)	(2.390)	(2.407)	(1.310)
College Students p.c.	29.873**	18.286**	29.737**	35.861***	18.928**	19.783**
	(39.545)	(25.927)	(39.431)	(47.942)	(27.261)	(28.765)
Share immigrants	0.021***	0.106*	0.021***	0.031**	0.103*	0.087**
	(0.030)	(0.142)	(0.030)	(0.043)	(0.138)	(0.102)
Share Aged 20-39	1.079***	1.074***	1.079***	1.072***	1.073***	1.071***
	(0.016)	(0.016)	(0.016)	(0.015)	(0.016)	(0.016)
Pop. share white	0.665	0.727	0.674	0.596	0.728	0.607
	(0.412)	(0.424)	(0.412)	(0.368)	(0.422)	(0.333)
Pop. share black	0.399	0.168**	0.389	0.371	0.166**	0.158***
	(0.279)	(0.121)	(0.280)	(0.259)	(0.125)	(0.099)
Center 1	0.389***	0.163***	0.388***	0.265***	0.158***	0.193***
	(0.100)	(0.042)	(0.101)	(0.084)	(0.040)	(0.056)
Center 2	3.787***	6.366***	3.858***	4.092***	6.550***	6.109***
	(0.914)	(1.496)	(1.080)	(0.829)	(1.633)	(1.409)
Election Closeness	0.642*	0.711	0.639*	0.699	0.709	0.654*
	(0.150)	(0.167)	(0.153)	(0.170)	(0.169)	(0.154)
Gri. × R: Org.		1.205***			1.198***	
		(0.045)			(0.058)	
Gri. × R: Network			1.007		1.012	
			(0.046)		(0.032)	
Gri. × R: Monetary				1.089*	1.011	
				(0.052)	(0.039)	
RMT						1.502***
						(0.107)
Gri. × RMT						1.143***
						(0.042)
Observations	3036	3036	3036	3036	3036	3036
Pseudo-R^2	0.749	0.756	0.749	0.751	0.756	0.753

Heteroskedasticity robust SE in parentheses; * p <0.1, ** p <0.05, *** p <0.01

Note: The exponentiated Poisson QMLE coefficients represent IRR for the model presented in Equation 2.2 for the year 2022. The RMT variable is the arithmetic mean of the three resource dimensions.

Table 2.8: Estimated effects (β-coefficients) of the sensitivity analysis.

	(1) OLS	(2) Neg. Bin.	(3) OLS int. M.	(4) Poi. int. M.	(5) Probit ext. M.
R: Organizations	0.282***	0.583***	0.522***	0.406***	0.243**
	(0.072)	(0.110)	(0.058)	(0.138)	(0.109)
R: Online Network	0.096***	0.192	0.171**	0.227**	0.185***
	(0.036)	(0.146)	(0.074)	(0.100)	(0.062)
R: Monetary Means	0.148***	0.338**	0.195***	0.375***	0.010
	(0.037)	(0.149)	(0.056)	(0.107)	(0.077)
R: Org. $\times$ R: Network	0.018	-0.159	-0.081	0.015	0.018
	(0.016)	(0.100)	(0.053)	(0.089)	(0.025)
Police Shooting B.	0.252***	-0.018	0.102**	0.019	0.320
	(0.073)	(0.085)	(0.048)	(0.051)	(0.205)
Unemployment	0.024***	-0.001	0.032**	0.085**	0.030*
	(0.008)	(0.047)	(0.016)	(0.043)	(0.017)
Gini	0.030***	0.010	0.037***	0.023	0.012
	(0.005)	(0.029)	(0.010)	(0.022)	(0.010)
Log Population	0.766***	1.424***	0.910***	1.027***	1.068***
	(0.040)	(0.110)	(0.058)	(0.078)	(0.068)
Density	-0.018	-0.013	0.009	0.023	-0.084**
	(0.027)	(0.082)	(0.035)	(0.050)	(0.041)
No school degree	-1.808***	2.818	-0.259	0.755	-3.735***
	(0.582)	(2.864)	(0.935)	(2.419)	(1.177)
College Students p.c.	2.837***	-0.812	-0.346	0.368	2.724**
	(0.761)	(2.291)	(0.954)	(1.553)	(1.326)
Share immigrants	3.589***	-1.007	1.027	-2.872**	4.496***
	(0.631)	(2.311)	(0.872)	(1.425)	(1.348)
Share Aged 20-39	0.020***	0.090***	0.049***	0.080***	0.003
	(0.006)	(0.025)	(0.010)	(0.015)	(0.013)
Pop. share white	-0.778***	-0.202	-0.783***	-0.139	-0.718*
	(0.192)	(0.597)	(0.263)	(0.586)	(0.424)
Pop. share black	-1.118***	-1.231	-1.196***	-1.065	-1.701***
	(0.202)	(0.767)	(0.323)	(0.727)	(0.443)
Center 1	1.473**	-0.719	-0.063	-0.887***	0.000
	(0.684)	(0.663)	(0.654)	(0.328)	(.)
Center 2	0.046	0.927	1.056*	1.197***	-1.907***
	(0.616)	(0.655)	(0.558)	(0.294)	(0.502)
Election Closeness	0.246**	1.119***	0.020	-0.568**	0.615***
	(0.096)	(0.371)	(0.120)	(0.230)	(0.169)
Observations	3036	3036	1403	1403	3032
(Pseudo-)R^2	0.685	0.187	0.642	0.688	0.427

Heteroskedasticity robust SE in parentheses; * p <0.1, ** p <0.05, *** p <0.01

Note: The table presents the sensitivity analysis for the model presented in Equation 2.1. The sensitivity analysis consists of the following alternative estimation approaches: specification (1) is an OLS estimation with the inverse hyperbolic transformed protest event count as the dependent variable; specification (2) is a negative binomial regression; specification (3) estimates the intensive margin based on OLS on the inverse hyperbolic transformed dependent variable; specification (4) is a Poisson QMLE estimation of the intensive margin; and specification (5) is a Probit estimation of the extensive margin (dropped variables due to perfect prediction of the outcome for *Center 1*).

Chapter 3

Rationalizing Protests – From Individuals to Networks

Abstract: We develop a theoretical model of network goods to explain protest participation. The model distinguishes between two equilibria: (i) when network effects are weak, individual participation is primarily driven by standalone benefits; (ii) when network effects are strong, expectations about others' participation become crucial. Using survey data from the American National Election Studies, we estimate a spatial autoregressive model and identify both standalone and network benefits through direct and indirect effects. Our findings reveal significant spillover effects in protest participation across individuals. Specifically, we identify both types of benefits for the full US sample as well as within more homogeneous subgroups. For most subgroups, protest dynamics are predominantly shaped by strong network effects. These estimation results, in combination with the model's implications, offer new insights into the scale of contemporary protest movements in the United States.[1]

Keywords: Collective Action; Public Protests; Network Effects; Social Networks.

[1] This research article is a collaborative effort between Jonathan Bothner, Maximilian Kähny, and me, with each author contributing equally to the paper. The research article is currently under review at the European Journal of Political Economy. The replication code can be accessed at `https://dx.doi.org/10.2139/ssrn.5018628` and the working paper version is available at `https://doi.org/10.7910/DVN/HDQFGL`.

3.1 Introduction

Public Protests have been a vital force for citizens to express their discontent in the public sphere worldwide. The act of demonstration – whether peaceful or violent – calls for policies, reforms, or even a change of government, and drives social and political change not only in democracies but also in authoritarian regimes. For instance, rallies that later became known as the Yellow Vest Protests arose in France from late 2018 until early 2019, when diesel prices increased by approximately 23% over 12 months, as the BBC (2018) reported.[2] With the ensuing riots drawing the participation of 282,000 protesters at its peak, the Yellow Vest Protests ultimately forced the French government to concede by announcing a halt on planned fuel tax increases and freezing electricity and gas prices for 2019. The economic damage of these civil outbursts amounted to 0.2 percentage points of quarterly growth – approximately 5 billion euros – according to Reuters (2019)[3], and was commented on by the Finance Minister of France, Bruno Le Maire, with these words:

> "It's a catastrophe for business, it's a catastrophe for our economy." —
> *BBC, 2018*

Considering the far-reaching implications of mass mobilization like the Yellow Vest Protests, the question arises: what precise dynamics unfold behind protests? We are particularly interested in understanding the mechanisms through which participation rates shift from very low to very high levels.

On the individual level, one rationale for participating in a protest is deriving protest-dependent utility by either satisfying a sense of moral duty or benefiting from an underlying objective. While both causes seem plausible, two issues related to the latter must be mentioned. First, most protests are directed toward the greater good of the public, leading to the classic free-rider problem among potential participants (Apolte, 2012; Olson, 1971). The second issue is related to the paradox of voting. It is present even if the free-rider problem is disregarded: With the individual's participation being non-pivotal in achieving the underlying objective of the protest, the cost of participation should outweigh the benefit of contribution (Downs, 1957). This cost is particularly substantial in the case of protests since, depending on the context or country, individuals might face not only minor transportation and opportunity costs, but severe repercussions like physical violence, incarceration, torture, and death.

[2] "Yellow vest protests 'economic catastrophe' for France", BBC, December 9, 2018, `https://www.bbc.com/news/world-europe-46499996`, last accessed: May 5, 2024.

[3] "French 'Yellow Vest' protests cost 0.2 percentage points of growth - Le Maire", Reuters, February 28, 2019, `https://www.reuters.com/article/idUSKCN1QH170/`, last accessed: May 5, 2024.

A more recent literature focuses on the role of networks in explaining individual participation decisions in protests. Similar to decisions regarding network goods, like joining a social media platform, an individual's value from protesting is assumed to increase as more people join (Economides, 1996). This network effect can be understood as direct and indirect benefits an individual gains from participating, which increase as the network attracts further participants. In the context of protests, network benefits include peer effects through social ties (Bursztyn *et al.*, 2021) and benefits caused by increased media coverage and larger protest venues (Enikolopov *et al.*, 2020). However, the realization of these network benefits crucially depends on others' participation. The individual's decision to participate is thus determined by their ex ante expectation about the protest. While Carter and Carter (2020) and Truex (2019) identify focal points as a mechanism to increase individual expectations about protest size, the general problem of coordination remains.

Combining these well-known insights – (1) the individual's utility from participating is subject to protest-dependent attitudes; (2) the individual's participation is non-pivotal to achieving the protest objective; (3) networks affect the individual's participation decision while allowing for multiple equilibria due to coordination problems – we put forward a single theoretical framework and empirically test its implications. In the framework, we consider a continuum of heterogeneous individuals whose utility from protesting consists of two components: (i) the interaction between an individual's type and the standalone benefit, and (ii) the network benefit that increases with the expected number of protest participants. Following the rational-expectation approach in the static model, we derive equilibrium outcomes for protest participation depending on the relation between the standalone and the network benefit. If the standalone benefit exceeds the network benefit, which we define as *weak network effects*, a unique equilibrium exists for any exogenous participation cost. If the network benefit exceeds the standalone benefit, which we define as *strong network effects*, two stable equilibria coexist for intermediate participation costs, namely the extreme outcomes of full and no participation. This result highlights the coordination problem individuals face when deciding to attend a protest, for which the network benefit is the determining factor for participation. In the dynamic version of the model, we further show the precise adjustment processes of protest participation over discrete time, which lead to the same outcomes in the limiting case as the static model.

Deducing hypotheses from the model dynamics and underlying intuitions, we use data on up to 6111 individual respondents from the American National Election Studies (ANES) 2020 (and 2016 for robustness) survey wave to test empirically (i) if the standalone and network benefits determine an individual's protest participation and (ii) the composition of both effects. The latter is particularly insightful for interpreting the model implications for protests in general, as well as the heterogeneities we explore by constructing more homogeneous societal groups.

The main empirical challenges consist of identifying standalone and network benefits and clustering individuals according to their political attitudes. Incorporating individuals' participation decisions within a social network into a simple ordinary least squares (OLS) regression leads to endogeneity due to the inherent interdependencies among them. Individual A's protest decision affects individual B's protest decision and vice versa. Hence, regressing individual A's on individual B's protest decision leads to simultaneity in the equation, resulting in biased and inconsistent estimates. To overcome this problem, we rely on a generalized spatial two-stage least squares estimator (GS2SLS) proposed by Drukker *et al.* (2013) in which the protest decisions of an individual's social network are instrumented by their individual characteristics. The standalone and network benefits are measured by interpreting the direct and indirect effects of a set of covariates. To elicit heterogeneities among different societal groups, we rely on the clustering algorithm proposed by Kaufman and Rousseeuw (2009) for political groups and a decision rule for more granular protest groups.

The results from the baseline regression support our first hypothesis since both the standalone and network benefits are found to be driving forces behind individual protest decisions. To explore heterogeneities, we divide individuals into three clusters based on political opinions. These clusters can be interpreted as groups of individuals within the sample holding left-leaning, moderate, and right-leaning political views. We find that the standalone and network benefits are significant determinants of the protesting decision within each cluster, though with varying magnitudes. Interestingly, network benefits seem to dominate standalone benefits for all clusters.

In a finer-grained specification, we pool individuals based on causes they consider important and identify four relevant potential protest movements: Black Lives Matter (BLM), healthcare, environment, and pro-immigration (these clusters are no longer mutually exclusive). Throughout various specifications, BLM, healthcare, and environmental protesting exhibit strong network effects, whereas no network effects are identified for immigration-related protesting. In the extension, we consider the model framework in combination with these benefit compositions to provide a novel explanation for their observed protest sizes.

This paper adds theoretical and empirical innovations to the literature. From a theory perspective, we argue that an individual's protest participation is non-pivotal to the success of the underlying objective and that the decision to protest is similar to consuming a network good. In this regard, our model strongly differs from the models of Acemoglu and Robinson (2001) and De Mesquita (2010), who assume protests to have a club good character. This assumption seems restrictive, as resulting policy changes are rarely targeted towards protest participants only – for example, the halt on a planned fuel tax rise achieved by the Yellow Vest Protests benefited not only the protesters but also all French citizens with an interest in low fuel prices. In contrast, our model allows for the public good character of protest and the associated free-rider problems. From an empirical point of view, we add to the literature by applying a spatial identification strategy based on building a synthetic network of US citizens. To the best of our knowledge, this approach to identifying network effects is novel to the protest literature.

The paper is structured as follows. After discussing the related literature in the following paragraphs, we introduce the framework in Section 3.2. In Subsection 3.2.1, we derive the potential equilibria of protest participation under strong and weak network effects in the static model. Next, we enrich the model by including discrete time in Section 3.2.2, deriving the precise protesting processes under weak and strong network effects. Thereafter, in Section 3.2.3, we formulate two hypotheses based on the model's implications. In Section 3.3, we introduce the empirical identification strategy used to test our hypotheses. The subsequent Section 3.4 presents the results for a baseline regression, as well as for political clusters in Subsection 3.4.2 and protest-specific clusters in Subsection 3.4.3. Finally, we conclude in Section 3.5. All proofs, additional figures, and tables are deferred to Appendix 3.6.1 and Appendix 3.6.2, respectively.

Related Literature Acemoglu and Robinson (2001, 2005) establish a first theoretical framework on the occurrence of protests by considering distribution battles among the population as the primary cause of protests and potential subsequent revolutions. Disadvantaged citizens utilize protests and the threat of revolution to pressure elites into creating a more equal distribution of wealth and income. These elites, who run the state, act under a revolution constraint that lets them strategically distribute resources. According to the authors, protests and revolutions occur when elites fail to adhere to the revolution constraint, and disadvantaged citizens respond by raising their voices via protests. However, some authors see an inconsistency between the explanation of Acemoglu and Robinson (2001, 2005) and fundamental insights from Olson's (1971) Theory of Collective Action (Egorov and Sonin, 2024; Apolte, 2012). In particular, Apolte (2012) claims that inequality among a nation's citizens is neither sufficient nor necessary for enlarged protests or revolutions. Instead, the ability to overcome the free-rider problem should explain protest events and revolutions.

The fundamental conflict between these two strands of theoretical literature emerges due to the different characterization of the underlying objective of protests: *club good* vis-à-vis *public good*. Under the notion of a club good, citizens protest to benefit from expected political changes (e.g., Acemoglu and Robinson, 2001, 2005). In contrast, under the assumption that the protest objective is characterized as a public good, the benefits from political changes are distributed among all citizens, regardless of the individual participation decision. Consequently, individuals would decide to free-ride – i.e., not protest – if the participation cost is sufficiently high. We do not impose any characterization on the nature of protests or the outcome. Instead, the micro foundation of our model is related to the literature on the voting paradox, which states that the individual's action is non-pivotal to the success of the underlying objective (Downs, 1957). This approach allows us to consider an individual's choice for protesting as the choice for a single network good, where protest participation is driven by a standalone and a network benefit. Thus, the conception of our model is also closely related to the industrial organization literature addressing network effects (Belleflamme and Peitz, 2015; Economides, 1996).

Later developments in the protest literature acknowledge how central the coordination problem is for explaining protests. This literature focuses on informational aspects, as, for example, Ellis and Fender (2011) extend the model from Acemoglu and Robinson (2001, 2005) by introducing information transmission via informational cascades based on a model proposed by Lohmann (1994). This informational exchange facilitates coordination. Another approach to overcome the coordination problem is proposed by De Mesquita (2010), who models a difference in the timing of movements. The model relies on a protest vanguard that moves first, thereby shaping the movement's trajectory. This move is observed by other citizens, who can subsequently join the protest. Both models incorporate informational exchange as a device for switching from a low to a high participation equilibrium. However, they maintain a club good characterization of protests, which makes them prone to the above-mentioned criticism, too. In our proposed model, we incorporate the presented informational aspects via expectation formation. Especially in the dynamic specification, informational cascades are modeled explicitly, and we provide an account of the pace at which they materialize.

In this regard, our paper is connected to the theoretical literature that centers on threshold models of collective behavior, as introduced by Granovetter (1978). In this framework, individuals choose to participate in collective actions – such as protests – when the number of prior participants exceeds a certain threshold. This idea provides a behavioral microfoundation for how participation cascades can emerge from heterogeneous individual preferences. Building on this, Wiedermann *et al.* (2020) present a network-based extension of Granovetter's model, showing how individual thresholds interact with social network structures to produce tipping points in protest dynamics. These models emphasize how even small shifts in expectations or participation can lead to rapid, large-scale mobilization, offering a complementary perspective to coordination-based and network benefit approaches.

Lastly, one recent strand of literature discusses the formation and effects of networks in overcoming the free-rider problem. The fundamental insight is that protest participation among social peers is strategically complementary. Such effects are identified by Bursztyn *et al.* (2021) in a field study with incentivized students in Hong Kong's anti-authoritarian protests and Enikolopov *et al.* (2020) using a natural experiment with social media data in Russia. Importantly, these approaches are capable of explaining protest participation irrespective of whether the desired outcome is a club good or public good. We aim to contribute to this literature by preserving the central role of networks in overcoming the free-rider problem while explicitly modeling the strategic complementarity and testing its resulting hypotheses with a spatial model. The spatial model is a promising technique for identifying network effects as highlighted by Souza (2024), and, to the best of our knowledge, has not been applied in the protest literature so far. Thereby, we provide complementary evidence in favor of Bursztyn *et al.* (2021) and Enikolopov *et al.* (2020) while also contributing to the question of whether protest participation is a strategic complement or substitute as discussed by Cantoni *et al.* (2019) and Shadmehr (2021).

3.2 The Framework

3.2.1 Static Model

To establish the baseline model for participating in demonstrations, we consider the textbook approach by Belleflamme and Peitz (2015) on markets for a single network good.[4] In the economy, suppose there is a continuum of heterogeneous individuals of mass 1 who derive utility from protesting with

$$U(\theta) = \theta a + vn^e - c\,, \tag{3.1}$$

where the individual's type θ is uniformly distributed on the interval $[0, 1]$. The utility of a representative individual to demonstrate consists of two components: (i) the interaction between an individual's type and the standalone benefit $a > 0$, which can represent but is not limited to a feeling of moral obligation, sympathy for the protest, or the benefit of the potential policy implementation to protest for; (ii) the network benefit $v > 0$ that,

[4] A related version of the model is put forward by Economides (1996).

for simplicity, increases linearly with the expected number of participants n^e. Note that this network benefit can encompass direct effects (social status, associating oneself with a movement) and indirect effects (larger venues, media representation). Participating in a demonstration comes at a non-monetary cost $c > 0$, which can be seen as opportunity costs and personal repercussions. Without loss of generality, assume that an individual's outside option $\underline{\mathrm{U}}$ is normalized to zero.[5]

For a given participation cost c and an expected network size n^e, the individual being indifferent between demonstrating or not is characterized by

$$\theta a + v n^e - c \geq \underline{\mathrm{U}}\,, \tag{3.2}$$

where

$$\theta \geq \frac{c - v n^e}{a} := \hat{\theta} \tag{3.3}$$

denotes the critical type threshold, i.e., individuals with types lower than $\hat{\theta}$ choose not to participate.

Ending up in a corner solution if $\hat{\theta}$ is not in the interval $[0, 1]$, the actual number of individuals demonstrating is

$$n(c, n^e) = \begin{cases} 0 & \text{if } n^e < (c-a)/v\,, \\ 1 - \frac{c - v n^e}{a} & \text{if } (c-a)/v \leq n^e \leq c/v\,, \\ 1 & \text{if } n^e > c/v\,. \end{cases} \tag{3.4}$$

We follow the rational-expectation approach to continue the analysis: if individuals form rational expectations in equilibrium, it must be that $n^e = n$.[6] From the demand function in (3.4), it is straightforward that a corner solution occurs if either $c > a$ (nobody participates with $n = n^e = 0$) or if $c < v$ (everyone participates with $n = n^e = 1$). It follows that both equilibria coexist if $v > a$. Furthermore, the interior equilibrium with $0 < n^e = n < 1$ is

$$n(c) = \frac{a - c}{a - v}\,. \tag{3.5}$$

First, suppose we are in the scenario where the standalone benefit exceeds the benefit

[5] In Appendix 3.6.1, we show that incorporating a non-zero outside option – which can reflect the free-rider problem – has no qualitative effect on the equilibrium outcome.

[6] The (Nash-)equilibria rely on self-fulfilling prophecies, where each individual has no incentive to change their action when rational expectations are satisfied.

of the network, $v < a$, which we define as *weak network effects*. In that case, the number of participants $n(c)$ is a decreasing function of the cost incurred. Thus, there is a unique participation level for any (exogenous) cost c: $n = 0$ for $c > a$, $n = (a - c)/(a - v)$ for $v \leq c \leq a$, and $n = 1$ for $c < v$.

Second, suppose we are in the scenario where the benefit of the network exceeds the standalone benefit, $v > a$, which we define as *strong network effects*. In that case, the number of participants is an increasing function of the cost incurred. As an intuition, the marginal willingness to incur costs increases with a marginal increase in participation numbers. Furthermore, there coexist three rational-expectations equilibria for intermediate costs ($a < c < v$): $n = 0$, $n = (c - a)/(v - a)$ and $n = 1$. By the argument of a dynamic adjustment process, i.e., a marginal cost increase (decrease) drives the equilibrium towards full (no) participation, we can disregard the unstable interior equilibrium (see Rohlfs, 1974). In contrast, the two remaining equilibria depict precisely the coordination problem individuals face when deciding on a protest for which the network benefit v is the determining participation factor. While both equilibria are stable, it is straightforward that by following the Pareto criterion, the no-participation equilibrium ($n = 0$) is Pareto-dominated: each individual would be better off coordinating to end up in the $n = 1$ equilibrium.

An example of the potential equilibria, i.e., the unique equilibrium under weak network effects and the multiple equilibria under strong network effects, for an exogenous cost $\hat{c}$ is depicted in Figure 3.1.

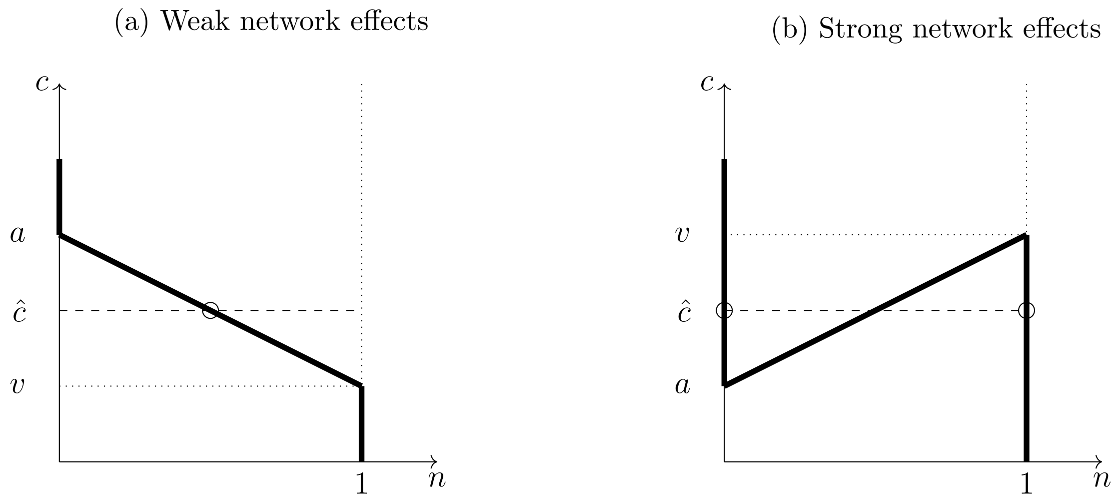


Figure 3.1: Equilibria under weak and strong network effects for given cost $\hat{c}$.

We can now characterize the following equilibria.

Proposition 1 (Static Network effects)**.** *In equilibrium, the number of individuals participating in the demonstration*

(i) under weak network effects ($v < a$) is denoted by

$$n(c) = \begin{cases} 0 & \textit{if}\, c > a\,, \\ \frac{a-c}{a-v} & \textit{if}\, v \leq c \leq a\,, \\ 1 & \textit{if}\, c < v\,; \end{cases} \tag{3.6}$$

(ii) under strong network effects ($v > a$) is denoted by

$$n(c) = \begin{cases} 0 & \textit{if}\, c > v\,, \\ 0 \veebar 1 & \textit{if}\, a \leq c \leq v\,, \\ 1 & \textit{if}\, c < a\,. \end{cases} \tag{3.7}$$

Conveying Proposition 1 to the real world, weak network effects may relate to protests about topics tied to goods (policies) and personal goals rather than social conditions and norms. Examples would be union wage talks, farmers' protests, or gun control, where the standalone benefit of the good, a, is the participant's priority, and higher private costs arguably lead to fewer participation numbers.

Examining the theoretical underpinnings of strong network effects in the context of protest participation, the BLM Movement, the 2020-2021 Belarusian Protests, and the Iranian Women's Rights Movement may be compelling examples. Despite the potential for high personal repercussions, these movements demonstrate how network effects can be powerful enough to motivate more individuals to participate (as long as the cost satisfies $c \leq v$). Furthermore, the coordination problem that arises under strong network effects, with its two stable corner equilibria, provides a theoretical foundation for the concept of focal points: despite no inherent structural change, there can be an equilibrium switch from no participation ($n = n^e = 0$) to full participation ($n = n^e = 1$) when instances, e.g., the killing of George Floyd, lead to higher expectations for protest.

3.2.2 Dynamic Model

To further analyze individuals' protest participation and its underlying processes, consider now the dynamic version of the model in discrete time $t \geq 1$, where the utility from demonstrating is given by

$$U_t(\theta) = \theta a + v n_t^e - c\,. \tag{3.8}$$

For simplicity, assume adaptive expectations where the expected number of participants at time t is the actual number of participants from the previous period with $n_t^e = n_{t-1}$, and that the initial participation level is exogenous with $n_0 \in [0,1]$. Holding the outside option fixed and normalized to $\underline{\mathrm{U}}_t = 0 \;\forall\; t \geq 1$, the critical type threshold at time t denotes

$$\theta = \frac{c - v n_{t-1}}{a} := \hat{\theta}_t \,, \tag{3.9}$$

where the number of participants is $n_t = 1 - \hat{\theta}_t$.

We can then show the following.[7]

Proposition 2 (Dynamic Network effects)**.** *The number of individuals participating in the demonstration at time $t \geq 1$ evolves according to*

$$n_t(c, n_0) = \begin{cases} 0 & \textit{if } n_o < \frac{c-a}{v}\,, \\ n_0 \left(\frac{v}{a}\right)^t + \left(1 - \frac{c}{a}\right) \sum\limits_{i=0}^{t-1} \left(\frac{v}{a}\right)^i & \textit{if } \frac{c-a}{v} \leq n_0 \leq \frac{c}{v}\,, \\ 1 & \textit{if } n_0 > \frac{c}{v}\,, \end{cases} \tag{3.10}$$

where in the interior, the change in the number of participants per period is defined by

$$n_t - n_{t-1} = \left(\frac{v}{a}\right)^{t-1} (n_1 - n_0) := \Delta_t \,. \tag{3.11}$$

Thus,

(i) under weak network effects ($v < a$), the interior solution converges to

$$\lim_{t \to \infty} n_t(c, n_0) = \frac{a - c}{a - v}\,; \tag{3.12}$$

(ii) under strong network effects ($v > a$), the interior solution converges to

$$\lim_{t \to \infty} n_t(c, n_0) = \begin{cases} 0 & \textit{if } n_0 < \frac{c-a}{v-a}\,, \\ 1 & \textit{if } n_0 > \frac{c-a}{v-a}\,. \end{cases} \tag{3.13}$$

Proposition 2 highlights the outcome of protest participation under weak and strong network effects, subject to the adjustment process in protest participation per period. Equation 3.10 states that protests, where the initial protest size is characterized by full (no) participation, remain at this corner solution. In contrast, protests, where the initial participation number satisfies the interval $(c - a)/v \leq n_0 \leq c/v$, follow the dynamic

[7] The proof is deferred to Appendix 3.6.1.

adjustment process in Equation 3.11: Under weak network effects ($v < a$), the change in protest participation, Δ_t, decreases per period, and the protest size ultimately converges to its interior limit $(a-c)/(a-v)$. With the standalone benefit being the determinant factor of protest participation, the marginal benefit of one more individual joining the protest on the critical type $\hat{\theta}_t$ exactly outweighs the individual's participation cost. Under strong network effects ($v > a$), the change in protest participation, Δ_t, increases per period. Here, the limit outcome of the protest size crucially depends on its initial participation number n_0, where Equation 3.13 shows that the protest size drives to the full (no) participation corner if $n_0 > (c-a)/(v-a)$ (if $n_0 < (c-a)/(v-a)$). The large (small) initial protest size in connection with the strong network benefit gradually increases (decreases) the critical type threshold $\hat{\theta}_t$, eventually leading to the corner solution.

This interplay between initial protest size and network effects may very well illustrate the workings of focal points. Focal points, when combined with strong network effects, can spur mass mobilization by attracting a sufficiently large number of initial protesters.

3.2.3 Model Implications and Hypotheses

Our theoretical model posits that individuals decide to participate in protests based on a utility function that includes both private (a) and social components (v). While the model does not yield sharp, testable implications, it provides a structure for identifying key mechanisms. We empirically test whether these mechanisms, standalone and network benefits, are significant predictors of participation. The following hypotheses guide our empirical analysis.

The first two hypotheses are direct premises of the model and address the cornerstone of our framework: the choice of a single network good can represent an individual's rationale for participating in a protest.

Hypothesis 1a. *The standalone benefit, as modeled in the utility function, increases an individual's protest participation.*

Hypothesis 1b. *The network benefit, as modeled in the utility function, increases an individual's protest participation.*

Turning to the implications of the model, the relationship between standalone and network benefits determines the equilibrium protest size. Under the assumption that participation costs are similar across different protests, strong network effects lead to extreme outcomes – that is, either small or large protests – whereas weak network effects lead to medium-sized protests.

Consistent with these insights, we expect substantial heterogeneity across societal subgroups in the effect sizes and relative importance of the standalone and network benefits. To test this, we formulate Hypothesis 2, which implies differences in the model's key parameters across different political and protest clusters. The second empirical challenge is thus to appropriately cluster individuals into more homogeneous groups and explore heterogeneities in the drivers of their protest participation.

Hypothesis 2. *The relative strength of standalone and network benefits differs across political and protest clusters.*

Potential heterogeneity implies that individuals' motivation for joining protests, and thus the relationship between standalone and network benefits, is protest-dependent. For instance, protests related to racial issues may have a different composition of standalone and network benefits than those related to healthcare policies. By identifying and comparing standalone with network benefits across different protests, this approach enables us to identify underlying strong or weak network effects.

3.3 Empirical Strategy

The empirical analysis tests the stylized model and the hypotheses derived in Section 3.2. Two primary empirical challenges arise. The first concerns the identification of the standalone benefit a and the network benefit v. This issue is addressed by employing a spatial autoregressive model, which allows for the estimation of both the standalone and network benefits in protest participation decisions. Both effects are assessed by estimating the direct and indirect effects of all employed covariates of the model. The standalone effect is represented by the direct effects, as these reflect changes in the individual's likelihood to protest due to changing personal characteristics. In contrast, the network benefit is captured by the indirect effects, which indicate feedback effects within an individual's network. The second challenge involves distinguishing between different political and protest clusters. Since individuals may align with multiple protest movements or political causes, assigning them to a single category may lead to misclassification. To address this, we analyze political and protest clusters separately, allowing for a more precise assessment of the model across individual movements. Political clusters are identified using the clustering algorithm proposed by Kaufman and Rousseeuw (2009), while protest clusters are constructed based on individuals' expressed concerns about societal issues in the United States.

In this section, we proceed as follows. First, we introduce the ANES database. The database choice is justified by the availability of required variables and their representativeness of the US population. Second, we argue that a spatial model is a good fit for collecting empirical evidence to test the proposed model. Next, we form political clusters and run the baseline model for each separately. Finally, we identify protest movement-specific potential protesters to compare the standalone and network effects across four distinct movements. This comparison allows us to assess the drivers of these movements and link them to their protest size.

3.3.1 Data

The ANES 2020 survey wave fulfills the crucial requirements for the selected empirical identification strategy. First, the database provides information on a representative sample from the US with extensive socio-demographic background information. More than 6,000 individuals[8] participated in the 2020 survey wave, which is a sufficiently large sample for investigating social networks based on individual characteristics. Most importantly, the database provides us with a combination of information about (1) protest behavior, (2) socio-demographic characteristics, and (3) political attitudes. Having this information in a representative sample of the US allows us to build an empirical model for individual protest decisions.

The empirical strategy relies on four categories of data included in the ANES database. Table 3.1 presents an overview of the utilized variables and their summary statistics. The binary protest variable indicates whether an individual respondent protested at least once during the past year. This binary protest variable serves as the dependent variable throughout this paper. Socio-demographic characteristics and personal attitudes are used as control variables and to calculate statistical proximity (network ties) between survey participants. The details of this procedure are explained below. The political attitudes category contains *thermometer* scores in which participants rate their feelings towards several societal groups (the higher the score, the friendlier the attitude). Some examples are attitudes towards conservatives, liberals, BLM protesters, and feminists. These attitudes form the basis for analyzing political clusters separately and comparing their respective standalone and network benefits.

To address potential concerns regarding the time frame of the primary analysis, we re-estimate the model using data from 2016 as a robustness check. Since survey respondents were asked between November 8, 2020, and January 4, 2021, whether they had participated in a protest within the past 12 months, all had the opportunity to protest during the summer, when COVID-19 restrictions were generally relaxed. However, while protest activity was legally permissible during the summer of 2020, some individuals may have

[8] After dropping those observations with missing values.

Variable	N	Mean	Std. Dev.	Min	Pctl. 25	Pctl. 75	Max
Protest Variable							
Protest	6111	0.10	0.30	0.00	0.00	0.00	1.00
Socio-Demographic Characteristics							
Age	6111	50.87	17.04	18.00	36.00	65.00	80.00
Education	6111	4.63	2.03	1.00	3.00	6.00	8.00
Income	6111	11.96	6.71	1.00	6.00	18.00	22.00
Black	6111	0.088	0.28	0.00	0.00	0.00	1.00
Male	6111	0.47	0.50	0.00	0.00	1.00	1.00
Health Status	6111	2.40	1.02	0.00	2.00	3.00	4.00
Married	6111	0.53	0.50	0.00	0.00	1.00	1.00
Children	6111	0.63	1.04	0.00	0.00	1.00	4.00
Personal Attitudes							
Left-right-scale	6111	5.56	2.77	0.00	4.00	8.00	10.00
Dev. Left-right-scale	6111	2.23	1.74	0.00	1.00	4.00	5.00
Religiosity	6111	2.88	1.50	1.00	1.00	4.00	5.00
Trust	6111	2.21	0.90	0.00	2.00	3.00	4.00
Longitude	6111	-91.62	15.63	-157.49	-97.33	-79.68	-69.76
Latitude	6111	37.80	5.01	21.11	34.61	41.31	61.41
Political Attitudes Towards...							
Labor Unions	6111	58.78	24.14	0	50	75	100
Conservatives	6111	53.51	27.81	0	40	70	100
Liberals	6111	51.03	29.33	0	30	70	100
Homosexuals	6111	59.42	26.88	0	50	85	100
Christian Fundamentalists	6111	45.78	28.97	0	25	60	100
Police	6111	69.99	25.41	0	60	85	100
Black Lives Matter	6111	54.07	35.56	0	15	85	100

Table 3.1: Summary Statistics. The data source is the American National Election Study Wave 2020; the list contains all variables employed in the empirical analysis. Longitude and latitude of respondents' locations are summarized at the state level.

remained hesitant to participate. To ensure that our findings are not driven by pandemic-related constraints, we replicate the baseline analysis for 2016. The model specification remains unchanged, with the exception of the religiosity variable, which is measured on a five-point scale in 2020 and as a binary indicator in 2016.

3.3.2 Identification Strategy

The empirical model is designed to disentangle standalone and network benefits in protest participation, as stated in Hypothesis 1a. To identify the standalone benefit – meaning the *utility* an individual receives from protesting, excluding any social network elements – we rely on *direct effects* of a set of predictor variables. Specifically, we assess the utility an individual receives merely because of protesting and expressing themselves without any social interaction before, during, or after the protest event. There are a variety of potential standalone protest determinants. To proxy individuals' intrinsic motivation to protest, we include their ideological distance from the political center on a left-right scale. The underlying assumption is that individuals with more extreme political views, i.e., those farther from the center, are ideologically more involved in the underlying protest objective and thus experience a larger intrinsic benefit. Expressive benefits are proxied by education, and psychological benefits by trust in others and religiosity. To include additional variation that is not fully proxied through this set of variables, we include general socio-demographic variables like gender, age, race, income, health status, marital status, and number of children. In sum, these are the covariates of the model for predicting individual participation and, importantly, their direct effects exclude any network-related utility as this is controlled for using the spatial term.

The identification strategy for the network benefit relies on a spatial autoregressive model. As discussed by, for example, Souza (2024), a spatial model is appropriate to identify network effects among individuals. That is, if individual A is a close friend of individual B, then A's decision to protest should influence B's decision and vice versa. The spillover is not limited to positive reinforcement. A's decision not to protest should consequently make individual B less likely to join the protest. Such spillover effects are identifiable via spatial regressions (Souza, 2024). Thus, including a spatial term in a regression analysis in combination with suitable individual-level covariates allows for the identification of the effects explored in the protest decision model from Section 3.2.

The estimated spatial autoregressive parameter, ρ, captures the degree of network-based interdependence in protest participation. While ρ itself is not directly interpretable as an individual-level effect, it reflects the equilibrium relationship among individuals in the network (LeSage, 2015). Therefore, we disaggregate the total effect of the covariates into direct and indirect effects. The direct effect of an increase in, for example, income is how much more likely the individual is to protest when the income increases by one unit, excluding feedback effects of the network. Once the individual is more likely to protest as a consequence of the income shock, their peers are also more likely to protest, and thus they have a second-order effect on the previously shocked individual. These higher-order effects are captured by the indirect effect, which we estimate for all covariates. In the SARAR (spatial autoregressive model with autoregressive disturbances) setting, where the dependent variable and disturbances are spatially lagged, the ratio of direct to indirect effects is the same for all covariates[9]. The ratio provides the crucial insight into how much an individual is affected by changing standalone benefits (direct effect) and how much by network benefits (indirect effects). In sum, the spatial autoregressive model allows for the disaggregation of the standalone and network benefits and, thereby, allows for estimating the parameters of the proposed model.

The first step in building spatial models is to grasp the social network structure within

[9] The size of the direct and indirect effects is determined by the weighting matrix, β coefficients, and ρ.

the sample. For this, we construct a quantitative measure of social proximity between individuals. We rely on the concept of homophily – meaning that individuals who share similarities are more attracted and exposed to each other and, therefore, are more likely to build friendships and partnerships (McPherson *et al.*, 2001). The first empirical challenge is thus to measure how similar each individual is to all other individuals in the sample. We rely on the inverse of the Mahalanobis distance, which measures how different two individuals are based on a set of variables. The variables used to construct the Mahalanobis distance include: a ten-item left-right political attitude scale, religiosity, trust in others, age, education, income, and the geographic coordinates (longitude and latitude) of the respondent's state of residence. Unlike the Euclidean distance, the Mahalanobis distance accounts for the variables' covariances and hence does not weigh highly correlated variables too strongly (Grimm and Yarnold, 2000). All variables are standardized prior to distance calculation.

Due to the sample size, we rely on statistical distances between individuals in our sample to construct a synthetic network for the entire dataset. A database containing just over 6,000 individuals across the US is, unsurprisingly, too small to identify actual social networks within the population. A solution is to rely on statistical distances among individuals. The distance between individuals A and B thus does not indicate how likely it is that both have a real network tie – e.g., a friendship – but rather indicates that individual A has individuals in their network that are similar to B. On this account, each individual within the sample has a synthetic network based on the whole ANES sample. This approach allows us to investigate whether an individual's synthetic network affects the individual's protest decision.

Equation 3.14 presents the regression model, including covariates, a spatial term, state fixed effects, and an error term. The dependent variable P_i is a binary variable equal to 1 if individual i participated in at least one protest event during the past 12 months and zero otherwise. The matrix X_i includes control variables for sex, marital status, children, age, race, income, education, health, extremism, religiosity, and trust. Details about these variables are available in Table 3.6 in the appendix. Spatial autocorrelation is assessed using the spatial term $W_{invdist}\ P_{-i}$. The first element is a spatial weighting matrix indicating the inverse distances between each individual i and all other individuals -i, and the second element denotes the other individuals' protest decisions P_{-i}. Because individuals are unlikely to be influenced by their whole network, we restrict spillover effects to their fifteen closest neighbors[10]. $\delta_{s(i)}$ denotes state fixed effects for all individuals and

[10]In the weighting matrix, the fifteen nearest neighbors are assigned a value of 1, while all others are set to 0. Then the weighting matrix is row normalized, a standard procedure in estimating spatial models (LeSage, 2015). The number of fifteen closest neighbors is varied in a sensitivity analysis to demonstrate

ϵ_i is the spatially correlated error term as depicted in Equation 3.15.

$$P_i = \alpha\, \mathbf{X}_i + \lambda\, \mathbf{W}_{invdist}\, P_{-i} + \delta_{s(i)} + \epsilon_i \tag{3.14}$$

$$\epsilon_i = \rho\, \mathbf{W}_{invdist}\, \epsilon_i + u_i \tag{3.15}$$

Simple OLS estimation of the model equation above would lead to biased estimates due to the simultaneity between the dependent variable from individual i and all other individuals -i. Therefore, we rely on the GS2SLS estimator proposed by Drukker *et al.* (2013). This estimator is based on spatial models derived earlier by Kelejian and Prucha (1998, 1999, 2010). The estimators instrument the simultaneous spatially lagged dependent variable with the remaining exogenous variables of each individual. Thereby, they allow for consistent estimation of spatial autoregressive models. To account for spatially correlated errors, we extend the SAR model to a SARAR specification, which explicitly allows for error terms to be correlated across the network. This ensures that our estimates remain robust in the presence of network-based error dependence.

The regression framework serves as the foundation for estimating both direct and indirect effects of each explanatory variable, which constitute the core of our identification strategy for disentangling standalone and network benefits. These effects are interpreted as follows: a one-unit increase in variable x leads to a predicted increase in outcome y of α_1 units directly and α_2 units indirectly. Here, α_1 captures the direct effect of a variable, while α_2 reflects the indirect effect, which aggregates all feedback effects transmitted through the network. As y increases due to the direct effect, the individual influences their peers, who in turn affect the original individual, generating recursive spillovers. Under a positive network parameter ρ, this process amplifies the initial shock. Accordingly, we distinguish between two components of the total effect: the direct effect, which we interpret as the standalone benefit, and the indirect effect, which we interpret as the network benefit. Importantly, both effects are necessarily interpreted as a consequence of the same increase in variable x.

Equations 3.16 to 3.18 define the direct, total, and indirect effects following LeSage (2015). These effects are derived from the spatial regression model and estimated to interpret the influence of explanatory variables in the presence of spatial dependence. Specifically, the direct effect, $\overline{M}(r)_{\text{direct}}$, captures the average impact of a change in a covariate on the outcome without taking the network into consideration. The total effect, $\overline{M}(r)_{\text{total}}$, measures the average overall impact of a change in a covariate for one indi-

the robustness of the results.

vidual on the outcome variable across all individuals in the network. The indirect effect, $\overline{M}(r)_{\text{indirect}}$, is then obtained as the difference between the total and direct effects and reflects the average spillover effect, i.e., how a change in one individual's covariate affects others in the network through spatial or social connections.

These effects are computed using the matrix $S_r(\mathbf{W}) = (I_n - \rho\mathbf{W})^{-1} I_n \alpha_r$, where I_n is the identity matrix of size n, ρ is the spatial autoregressive parameter, $\mathbf{W}$ is the spatial weighting matrix, and α_r is the coefficient of the r-th covariate. The term $(I_n - \rho\mathbf{W})^{-1}$ captures the spatial multiplier effect, indicating spillover effects throughout the network. The trace (tr) operator in the direct effect formula averages the diagonal elements of the impact matrix, while the total effect is computed as the average of all elements (l_n being a $n \times 1$ vector of ones). The spatial regression, the decomposition into direct and indirect effects, and the associated test statistics are estimated using the *spatialreg* R package provided by Bivand *et al.* (2021).

$$\overline{M}(r)_{\text{direct}} = n^{-1} \operatorname{tr}(S_r(\mathbf{W})) \tag{3.16}$$

$$\overline{M}(r)_{\text{total}} = n^{-1} l_n' S_r(\mathbf{W}) l_n \tag{3.17}$$

$$\overline{M}(r)_{\text{indirect}} = \overline{M}(r)_{\text{total}} - \overline{M}(r)_{\text{direct}} \tag{3.18}$$

where

$$S_r(\mathbf{W}) = (I_n - \rho\mathbf{W})^{-1} I_n \alpha_r \tag{3.19}$$

3.3.3 Identifying Political Clusters

After having obtained average results across a representative sample of the US population, we are interested in exploring heterogeneities among different societal groups across the US. Considering the model dynamics, these differences in the composition of benefits directly translate into different equilibrium outcomes for protests. The key objective of forming groups across the US population is thus to explore potential differences in the causes of their protest participation. To do so, we build clusters and assess the standalone and network benefits for them separately. There are two subsets we investigate. First, different individuals are clustered into political groups, and we examine whether the protest participation determinants (i.e., standalone and network benefits) vary across these political groups. Second, individuals are clustered based on their likelihood of participating in different protest movements, and then we assess whether the protest determinants vary across these clusters. Both specifications contribute to testing Hypothesis 2 as they explore heterogeneities among protest determinants that help explain differences in protest sizes across movements.

The first heterogeneity analysis consists of differentiating protests according to their *political directions*. We aim to elicit differences between political groups in terms of the underlying reasons for their protest decisions – i.e., are there substantial differences in the standalone and network benefits across different political clusters? To construct political clusters, we rely on a clustering algorithm by Kaufman and Rousseeuw (2009). This algorithm is based on several protest-sensitive political attitudes and forms clusters according to similarity between individuals. We use this algorithm to minimize subjective assumptions in the clustering process.

Our approach relies on survey respondents' attitudes towards a set of stylized groups of people within US society. These groups are listed in Table 3.1 under "Political Attitudes Towards...". These variables are scores from 0 (very low opinion) to 100 (very high opinion) for the following groups: labor unions, conservatives, liberals, feminists, Christian fundamentalists, police, and BLM participants. Clustering individuals according to their attitudes towards these groups should result in groups that share concerns about similar political issues and hold relatively similar political views. For example, individuals with a very low opinion of feminists and a high opinion of conservatives and Christian fundamentalists are more likely to attend the same protest movements. The same applies to individuals with opposing views, who are likely to participate in different types of protest events.

The individuals are clustered according to the partitioning around medoids (PAM) algorithm by Kaufman and Rousseeuw (2009). The authors built the K-medoids algorithm, which is more robust to outliers than the more common K-Means algorithm (Kassambara, 2017). A medoid is an observation that is representative of a whole cluster, and all other observations sort themselves around k-medoids. The algorithm determines the choice of medoids; the clustering algorithm works as follows (Kaufman and Rousseeuw, 2009).

1. Randomly select k individuals from the sample. These are the first cluster-medoids.
2. Calculate the Euclidean dissimilarity matrix for all individuals based on the variables "Political Attitudes Towards..." from Table 3.1.
3. Assign each individual to the cluster-medoid to which it is most similar.
4. Within each cluster, check if there is an alternative medoid that minimizes the average dissimilarity coefficient of the cluster. If this is the case, this alternative medoid is the new cluster-medoid.
5. In case one or more medoids change, return to (3); otherwise, end the algorithm.

Since the ANES database contains more than 6,000 individuals, the PAM algorithm is computationally too expensive to estimate. Therefore, we rely on the Clustering Large Applications (CLARA) extension to the k-medoids algorithm, which is also based on Kaufman and Rousseeuw (2009). The CLARA algorithm splits the sample into multiple subsets – 50 in our case. For each subgroup, the k medoids are determined using the PAM algorithm. Thus, we end up with 50 sets of potential medoid combinations. All sets are assessed using the full dataset. The remaining observations are assigned to whichever medoid is closest to them, separately for each set of medoids. Then, the average dissimilarity between the observations and their respective medoid is calculated. The set of medoids that minimizes this number is selected.

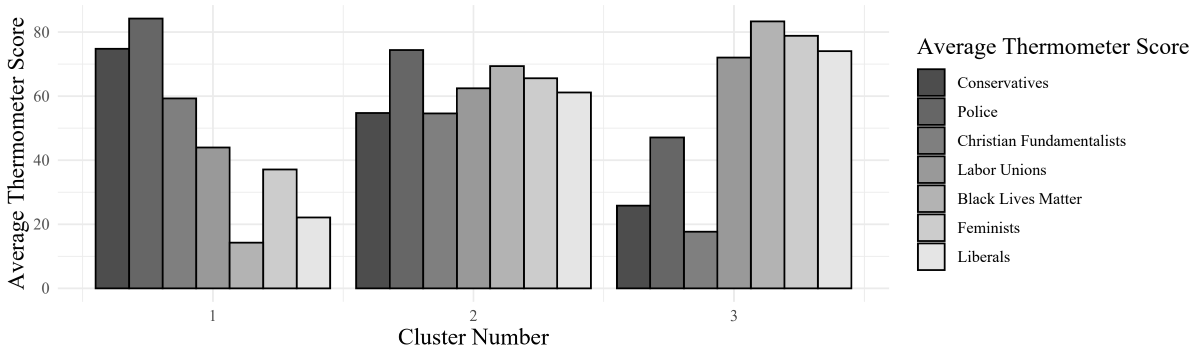


Figure 3.2: Average Thermometer Scores for Political Clusters. Cluster 1 includes 2,455 individuals, cluster 2 includes 2,520 individuals, and cluster 3 includes 1,897 individuals.

Applying the CLARA algorithm based on the seven political attitudes, we identify three clusters as illustrated in Figure 3.2. Individuals in cluster 1 have, on average, relatively high opinions about conservatives, the police, and Christian fundamentalists. This cluster thus seems to include individuals who align themselves with conservative ideologies in the US. The average scores for cluster 3 are roughly the opposite. Individuals in cluster 3 have, on average, relatively high opinions about liberals, feminists, BLM protesters, and labor unions. Thus, this cluster seems to include left-leaning individuals. Cluster 2 seems to collect all individuals who do not fit into both political camps. Therefore, this cluster contains individuals who deviate atypically in their opinions from left- or right-leaning individuals or score each group equally well on the variables. We call this cluster the "moderates" as their attitudes toward the different groups are less extreme on average.

These cluster interpretations are helpful for assessing heterogeneity in standalone and network benefits. Assigning individuals to different clusters enables separate assessment of these effects. This approach aims to shed light on the drivers of left-wing versus right-wing protests and whether protest motivations differ between left- and right-leaning individuals. We re-estimate the model separately for each of the three clusters, using the same specification as in Equation 3.14.

3.3.4 Identifying Protest Clusters

The final empirical specification explores further heterogeneities by disaggregating the analysis to specific *protest movements*. This approach is intended to shed further light on Hypothesis 2, which claims that there are dominant network effects in (at least some) protest movements. To further identify the conditions under which network effects are most pronounced, we construct clusters of individuals based on the protest causes they support. This allows us to examine how the relative importance of the network and standalone benefits varies across these groups. Importantly, individuals may belong to multiple groups, as participation in one protest does not preclude involvement in another. Therefore, the clusters in this specification are not mutually exclusive.

The first step in building protest clusters involves assessing participants' awareness of political issues. The survey includes a question that asks participants to identify what they perceive as the most pressing problems facing the US. Respondents could name multiple issues. Based on respondents' answers, we identified four topics that were mentioned most frequently: healthcare (42.7%), race relations (19.3%), environment (6.6%), and immigration (4.3%).

Relying solely on these categories may be problematic because some topics do not directly translate to policy implications. For example, individuals skeptical of BLM protests can name race relations as the biggest problem in the US today, but pro-BLM individuals may also name the same problem. Both groups, however, derive very different policy conclusions from the same issue and are, therefore, very unlikely to join the same protest or to belong to the same protest movement's group of potential protesters. Thus, we combine the naming of the problem with the thermometer scores about political groups. From the group that judges race relations to be one of the most significant problems in the US, only those who indicated a thermometer score of 50 or higher toward the group "BLM protesters" remain in the group. All other individuals are excluded. Thus, this group represents people who see race relations as one of the most pressing issues in the country, while being in line with BLM demands. The individuals are thereby grouped according to their problem awareness and their attitudes toward these protest movements. Based on individuals' survey responses, we thus assign 396 individuals to the group "Environment", 877 individuals to the group "Black Lives Matter", 260 individuals to the group "Immigration", and 2,580 individuals to the group "Healthcare". We consider an individual who is a member of one of these groups to be a potential protester for the corresponding issue.

We estimate the model specified in Equation 3.14 separately for four subsamples, each corresponding to a distinct protest movement. This second extension enables us to examine whether standalone and network benefits vary across movements. A key advantage of this approach is that individuals may belong to multiple groups, or none at all, depending on their political attitudes. This allows us to define potential protesters with precision and to quantify the magnitude of both standalone and network benefits within this subset. In a subsequent analysis, we relate the ranking of these estimated effects to the observed levels of participation across the four protest movements.

3.4 Results

In the following section, the theory-deduced hypotheses are tested by relying on the presented empirical identification strategy. As a first step, Hypotheses 1a and 1b are tested for all individuals across the US sample. The hypotheses indicate that standalone and network benefits shape individual protest participation. Subsections 3.4.2 and 3.4.3 explore heterogeneities among different societal groups regarding the benefit composition. Thereby, these specifications shed light on Hypothesis 2, which indicates that given the stark contrast of aggregate protest outcomes, we expect substantially different benefit compositions between different protest movements.

3.4.1 Baseline Estimation

The baseline estimation follows the identification strategy described in Subsection 3.3.2, where a spatial weighting matrix is defined via which each individual is affected by their peers. To demonstrate robustness to methodological choices, we present the results for different specifications in which spillovers are given by the 10, 15, or 20 nearest neighbors. Furthermore, we include a set of covariates commonly applied in the literature to estimate their effects on an individual's likelihood to participate in protests. The model is estimated using the presented GS2SLS estimator.

Table 3.2 presents the spatial regression output for the baseline regression, indicating the presence of robust network effects and the significance of the included covariates. Importantly, the regression coefficients are not yet interpretable because a shock to a particular individual has an impact on the spatial (network) equilibrium, which in turn has feedback effects (LeSage, 2015). We can merely state that some of the covariates, like age, education, or marital status, are likely to affect individual protest participation. As we observe a statistically significant λ and ρ, we conclude that there is indeed substantial spatial (network) dependence within the analyzed sample. Hence, OLS estimation of the same model is biased and inconsistent, while the SARAR specification effectively accounts for spillover effects across the network.

Once we obtain the spatial regression output, we estimate the direct, indirect, and

Table 3.2: Baseline Spatial Regression Results.

	Model (1) Neighbors: 10	Model (2) Neighbors: 15	Model (3) Neighbors: 20
Intercept	0.168***	0.154***	0.146***
	(0.049)	(0.046)	(0.048)
Male	-0.002	-0.002	-0.002
	(0.007)	(0.007)	(0.007)
Age	-0.003**	-0.003**	-0.003**
	(0.001)	(0.001)	(0.001)
Age^2	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Black	0.017	0.017	0.018
	(0.014)	(0.013)	(0.014)
Income	0.000	0.000	0.000
	(0.001)	(0.000)	(0.001)
Education	0.004**	0.003*	0.003
	(0.002)	(0.002)	(0.002)
Left Right Scale	-0.012***	-0.011***	-0.010***
	(0.003)	(0.002)	(0.002)
Left Right Dev.	0.018***	0.017***	0.017***
	(0.003)	(0.003)	(0.003)
Health	0.003	0.002	0.002
	(0.004)	(0.004)	(0.004)
Married	-0.020**	-0.021**	-0.021**
	(0.008)	(0.008)	(0.008)
Number of children	-0.011***	-0.011***	-0.011***
	(0.004)	(0.003)	(0.004)
Religious	0.000	-0.001	-0.001
	(0.002)	(0.002)	(0.002)
Trust government	0.005*	0.005*	0.005*
	(0.003)	(0.003)	(0.003)
North East	-0.006	-0.005	-0.004
	(0.010)	(0.009)	(0.010)
Mid West	-0.017*	-0.014	-0.013
	(0.009)	(0.009)	(0.009)
South	-0.021**	-0.019**	-0.018**
	(0.009)	(0.008)	(0.009)
Spatial Lag Y: λ	0.569***	0.637***	0.676***
	(0.124)	(0.117)	(0.126)
Spatial Lag ϵ: ρ	-0.504***	-0.641***	-0.555***
	(0.107)	(0.114)	(0.140)
Observations	6111	6111	6111

Heteroskedasticity robust spatial SE in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Baseline SARAR regression results based on Equation 3.14. The full ANES sample is used. Columns differ by the number of nearest neighbors used in the spatial weighting matrix (10, 15, and 20, respectively).

total effects of each covariate based on the spatial parameter λ, the weighting matrix $\mathbf{W}$, and the variable's spatial point estimate α. The results of this estimation are presented in Table 3.3. As can be seen in the table, the direct effects are consistently statistically identified, while for the indirect effects, only the variable *Left Right Scale* remains a statistically significant predictor for individual protest participation. A one-unit increase in *Left Right Scale* (equivalent to becoming "more right") decreases one's likelihood to participate in a protest by 1.1 percentage points directly and by 1.9 percentage points indirectly, amounting to a total effect of 3.0 percentage points. The indirect effect is merely driven by the individual's effect on the network equilibrium, where the individual is influenced by the network's feedback effects. A more substantial total effect is observed for the deviation from the political center. If *Left Right Deviation* increases by one unit, the individual is, altogether, 4.7 percentage points more likely to join a protest event.

Table 3.3: Baseline Impacts.

Effect	Direct	Indirect	Total
Male	-0.002	-0.003	-0.004
	(0.007)	(0.023)	(0.028)
Age	-0.003**	-0.005	-0.008*
	(0.001)	(0.004)	(0.005)
Age^2	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Black	0.018	0.030	0.048
	(0.014)	(0.051)	(0.060)
Income	0.000	0.001	0.001
	(0.001)	(0.002)	(0.002)
Education	0.003*	0.005	0.008*
	(0.002)	(0.003)	(0.004)
Left Right Scale	-0.011***	-0.019*	-0.030***
	(0.002)	(0.011)	(0.010)
Left Right Dev.	0.017***	0.029	0.047**
	(0.003)	(0.023)	(0.022)
Health	0.002	0.003	0.006
	(0.004)	(0.011)	(0.014)
Married	-0.021***	-0.036	-0.057
	(0.008)	(0.038)	(0.042)
Number of children	-0.011***	-0.019	-0.030
	(0.003)	(0.021)	(0.022)
Religious	-0.001	-0.001	-0.002
	(0.002)	(0.006)	(0.007)
Trust government	0.005*	0.008	0.013
	(0.003)	(0.012)	(0.014)
North East	-0.005	-0.008	-0.013
	(0.009)	(0.029)	(0.036)
Mid West	-0.015	-0.025	-0.039
	(0.009)	(0.020)	(0.026)
South	-0.019**	-0.033	-0.052*
	(0.008)	(0.030)	(0.032)

Heteroskedasticity robust spatial SE in parentheses; * p <0.1, ** p <0.05, *** p <0.01

Note: Impact estimation of baseline SARAR regression results based on Equation 3.14. The full ANES sample is used. Columns differ by the estimated effect type. Results are obtained for the specification that includes the 15 nearest neighbors in the weighting matrix.

Interpreting these results in light of the theory-deduced hypotheses, we find evidence for Hypotheses 1a and 1b as both standalone (direct effects) and network (indirect effects) benefits affect protest participation. Neither of the effects seems to dominate the other in size, leaving us uncertain in which of the model-derived equilibria we are located (Proposition 1). If standalone benefits dominated, we would assume a unique equilibrium solution with mediocre participation, while in the opposite case, we would assume either no or full participation as an equilibrium solution.

The findings connect to a broader literature by indicating that protests are games of strategic complements. This fact becomes clear when interpreting the consistently positive statistical identification of λ, indicating positive autocorrelation within the network. The paper thus confirms findings from experimental studies about positive network effects in field experiments (Bursztyn *et al.*, 2021), quasi-experimental studies (Enikolopov *et al.*, 2020), and studies about online networks (Larson *et al.*, 2019; Mundt *et al.*, 2018). The contribution to these existing studies is the novel application of constructing a synthetic network and applying a spatial autoregressive model to estimate network dependence.

3.4.2 Political Clusters

In the previous section, we find evidence for Hypotheses 1a and 1b as both the standalone and network benefits are found to increase an individual's protest participation, as predicted by the proposed model. However, no clear dominance of either the standalone or network benefit emerges in the baseline regression. Hence, we cannot make any statements about which of the equilibria materializes in the case of differentiation in Proposition 1. This may change once we collect more fine-grained evidence in more homogeneous groups. To explore such heterogeneities, we apply the same estimation techniques to political clusters built according to the algorithm presented in Subsection 3.3.3. Within these clusters, individuals show much stronger homogeneity as they are sorted according to their political attitudes.

Table 3.4 presents the estimated spatial coefficients and the ratio of indirect to direct effects in the estimation of the political clusters. The parameter λ should, by definition, be an element between -1 and 1. As becomes visible, for example, for the Right-Wing specification with 15 neighbors, the point estimate exceeds this range. Considering the standard errors, the 95% confidence interval clearly includes values below 1, and hence the exceeding value is likely due to statistical imprecision. Therefore, we interpret the parameter nonetheless. However, calculating the indirect to direct effect ratio is no longer informative as the Leontief Expansion is only applicable when ρ is an element between -1 and 1.[11]

Table 3.4 provides consistent evidence of strong network dependencies. Across all

[11] If $|\rho| < 1$, the series of feedback loops can be simplified: $(I_n - \rho W)^{-1} = I_n + \rho W + \rho^2 W^2 + \rho^3 W^3 + ...$ (LeSage, 2015, p. 34). Once $|\rho| \geq 1$, this transformation is no longer possible, and, hence, we omit displaying the indirect to direct effect ratio.

specifications, the estimated spatial autoregressive parameter (λ) exceeds 0.8, indicating substantial spillover effects among individuals within political clusters. These findings suggest that network benefits outweigh the standalone benefits of individual characteristics. To assess the relative importance of these two channels, we compute the ratio of indirect to direct effects. In the SARAR specification, this ratio is the same for all covariates, allowing for a straightforward comparison. The resulting estimates imply that network effects are significantly larger than direct effects, but for some cases, the computation of effects is not feasible due to $\rho > 1$.

Table 3.4: Political Cluster Impacts

Political Cluster	Right-Wing	Moderates	Left-Wing
	Neighbors: 10		
Spatial Lag Y: λ	0.903***	0.848***	0.837***
	(0.180)	(0.118)	(0.271)
Spatial Lag ϵ: ρ	-0.612***	-0.773***	-0.409**
	(0.144)	(0.147)	(0.206)
Indirect effect/direct effect	8.24	5.01	4.66
	Neighbors: 15		
Spatial Lag Y: λ	1.066***	0.979***	0.789***
	(0.168)	(0.131)	(0.297)
Spatial Lag ϵ: ρ	-0.884***	-0.866***	-0.556**
	(0.168)	(0.182)	(0.242)
Indirect effect/direct effect	*omitted*	25.33	11.15
	Neighbors: 20		
Spatial Lag Y: λ	1.058***	0.950***	1.049***
	(0.178)	(0.137)	(0.275)
Spatial Lag ϵ: ρ	-0.807***	-0.900***	-0.738***
	(0.217)	(0.207)	(0.271)
Indirect effect/direct effect	*omitted*	17.33	*omitted*

Heteroskedasticity robust spatial SE in parentheses; * p <0.1, ** p <0.05, *** p <0.01

Note: Political cluster SARAR regression results based on Equation 3.14. Each column corresponds to a distinct political cluster. Covariate coefficients are omitted for brevity.

Applying this evidence to the proposed model, we argue that *within* political clusters, there are strong network effects. The resulting model prediction is that for these groups, there are equilibria with zero and with full participation. Hence, extreme cases of mobilization are predicted to prevail in protests within these groups. To get an idea of whether this is actually the case, we analyze even more homogeneous groups in the following subsection, where we specifically shed light on four specific protest movements.

3.4.3 Protest Clusters

While political clusters already help investigate network effects in more homogeneous groups of individuals, the objective of this subsection is to assign individuals to protest movement clusters. The advantage of doing so is that we can compare the measured effects to the aggregate protest outcome during 2020 and thereby have a preliminary test for the model's implications. To do so, four topics are chosen: BLM, immigration, health, and environment, which have been particularly salient in 2020. The detailed procedure is described in Subsection 3.3.4.

Interpreting the results in Table 3.5, we conclude that BLM, health, and environment protests show strong network effects while immigration-related protests show weak network effects. These results are driven by the high λ-parameter for the former three protest movements, which subsequently leads to high ratios of indirect to direct effects. For the immigration cluster, however, not even the spatial parameter λ is statistically identified. Given this lack of statistical identification, we cannot reject the null hypothesis that there is no spatial dependence. As a result, we predict the aggregate protest outcome of BLM, health, and environment to either show very high or very low participation as described in the case differentiation *strong network effects* in Proposition 1. In contrast, we expect immigration to show mediocre aggregate mobilization under the *weak network effects* equilibrium in Proposition 1.

Table 3.5: Protest Cluster Impacts

Protest Cluster	BLM	Immigration	Health	Environment
	Neighbors: 10			
Spatial Lag Y: λ	0.852***	-0.310	0.703***	0.896***
	(0.181)	(0.367)	(0.196)	(0.216)
Spatial Lag ϵ: ρ	-0.754***	0.335	-0.533**	-0.900***
	(0.239)	(0.352)	(0.176)	(0.291)
Indirect effect/direct effect	5.154	-0.241	2.204	7.656
	Neighbors: 15			
Spatial Lag Y: λ	0.850***	0.111	0.728***	0.968***
	(0.182)	(0.531)	(0.217)	(0.247)
Spatial Lag ϵ: ρ	-0.881***	-0.130	-0.548**	-0.900**
	(0.271)	(0.567)	(0.218)	(0.351)
Indirect effect/direct effect	5.281	0.125	2.545	26.017
	Neighbors: 20			
Spatial Lag Y: λ	0.987***	0.066	0.941***	1.014***
	(0.198)	(0.693)	(0.202)	(0.307)
Spatial Lag ϵ: ρ	-0.900***	-0.262	-0.885***	-0.900**
	(0.316)	(0.718)	(0.232)	(0.405)
Indirect effect/direct effect	65.384	0.070	14.780	*omitted*
Observations	885	264	2605	401

Heteroskedasticity robust spatial SE in parentheses; * p <0.1, ** p <0.05, *** p <0.01

Note: Protest cluster SARAR regression results based on Equation 3.14. Each column corresponds to a distinct protest cluster. Covariate coefficients are omitted for brevity.

When considering the aggregate protest outcomes in Figure 3.3, the previously mentioned predictions are consistent with these outcomes. (Anti-)Racism, as a protest topic, is by far the most salient protest topic during 2020. From the modeling perspective, individuals managed to coordinate into the high participation equilibrium for this category of protests. By contrast, protests related to environmental and health issues remain sparse and small in scale, suggesting that despite the presence of strong network effects, individuals failed to coordinate and thus remained in a low-participation equilibrium. Finally, immigration-related protests are mediocre in number and size, which is also in line with the model's implications.

The connection of the individual-level findings to the aggregate protest outcome should be understood as an illustrative analysis rather than a statistical test, as there are some limitations to this endeavor. First, there is no measure of the individual participation cost for either of the protests. Partly, these costs are proxied by the covariates, and generally, they may not differ too much, given the democratic right to protest in the US. Furthermore, we acknowledge the limited number of protest categories and a certain arbitrariness in judging immigration to *already* be of mediocre size. However, the varying observational units and different participation rates within the different clusters are not a limitation of this analysis. Spillovers are measured in both positive and negative directions, and hence, there can be strong network effects even if only a few individuals protest in the cluster. Also, cluster size should have no impact on the coefficient size; if anything, small clusters may show larger standard errors due to lower statistical power.

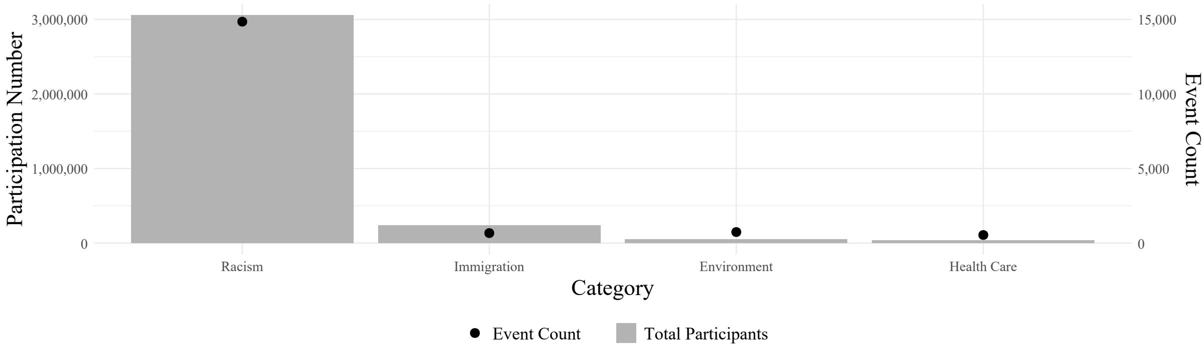


Figure 3.3: Illustration of protest size and frequency grouped by protest issues for 2020 according to the Crowd Counting Consortium database.

As discussed previously, we find strong evidence for the strategic complementarity of protests for BLM, health, and the environment. Protest mobilization dynamics for these categories can thus be modeled as coordination games in which both the high and low participation equilibria may materialize, similar to De Mesquita (2010). The murder of George Floyd on May 25, 2020, sparked large protests, shifting participation from low to high. In the context of coordination games, this event qualifies as a Focal Point, which makes a certain equilibrium salient and thus leads to coordination. Such mechanisms are first described by Schelling (1960) from a game-theoretical perspective. Other scholars investigate the role of focal points, mostly interpreted as particular days associated with a country's culture and history, in protest coordination (Truex, 2019; Carter and Carter, 2020; Ketchley and Barrie, 2020). For environmental and healthcare-related protests, no sufficiently salient event occurred to trigger large-scale protests.

3.5 Conclusion

This paper provides a theoretical model of protesting and empirically tests the model's implications. First, we apply a novel framework for individuals' protest decisions by adopting a model of a single network good by Belleflamme and Peitz (2015). In contrast to standard literature, we argue that the individual's protest participation is non-pivotal to the success of the underlying protest objective and that the utility derived from protesting consists of two components: (i) the interaction between an individual's type and the standalone benefit, and (ii) the network benefit which increases linearly with the expected number of further participants. We then derive equilibrium outcomes where the relation between the standalone benefit and the network benefit determines the aggregate outcome: Under *weak network effects*, a unique equilibrium exists for any exogenous participation cost. Under *strong network effects*, two stable equilibria coexist for intermediate participation costs, namely the extreme outcomes of full and no participation, which illustrates the coordination problem of protests.

The results from the empirical investigation speak in favor of the model dynamics. Using a spatial auto-regressive model, we identify the standalone and network benefits as drivers of individual protest participation in the whole US sample. We compare the estimated indirect and direct effects of the model's covariates to draw conclusions about whether the protesting decision is characterized by strong or weak network effects. We find evidence for strong network effects in three clusters of individuals with similar political attitudes. Furthermore, our results suggest heterogeneity in the strength of network effects across different protest movements. Protests related to race relations, healthcare, and the environment appear to be characterized by strong network effects. Pro-immigration protests, on the other hand, are driven by weak network effects. This heterogeneity in the relative sizes of the standalone and the network benefit across different groups of potential protesters provides a novel explanation for the differences in protest sizes and mobilization success of actual protests.

3.6 Appendix

3.6.1 Mathematical Appendix

Outside Option $\underline{U}$. First, suppose that an individual's outside option is type-dependent with

$$\underline{\mathrm{U}} = \theta \underline{\mathrm{u}}\,, \tag{3.20}$$

where θ is uniformly distributed on the interval $[0, 1]$. Then, the critical type threshold denotes

$$\hat{\theta} := \frac{c - vn^e}{a - \underline{\mathrm{u}}}\,, \tag{3.21}$$

where it is straightforward to see that the type-dependent outside option only affects the level and slope of the function $n(c)$.

Second, suppose that an individual's outside option scales with the expected number of participants, i.e.,

$$\underline{\mathrm{U}} = \phi n^e \tag{3.22}$$

with base utility $\phi > 0$. Then, the critical type threshold denotes

$$\hat{\theta} := \frac{c - (v - \phi)n^e}{a}\,, \tag{3.23}$$

where we can readily see that the scaling outside option, ϕn^e, only affects the steepness of the function $n(c)$.

With the above cases being variable transformations, it is without loss of generality to normalize an individual's outside option to $\underline{\mathrm{U}} = 0$. □

Proof of Proposition 2. Recall the utility function $U_t(\theta) = \theta a + vn_t^e$, where $n_t^e = n_{t-1}$ and the initial participation level $n_0 \in [0, 1]$. Recall further that with an outside option $\underline{\mathrm{U}}_t = 0$, the critical type threshold at time t denotes $\hat{\theta}_t = (c - vn_{t-1})/a$.

At $t = 1$, the number of participants at a protest denotes

$$n_1(c, n_0) = \begin{cases} 0 & \text{if } n_0 < \frac{c-a}{v}\,, \\ 1 - \frac{c - vn_0}{a} & \text{if } \frac{c-a}{v} \leq n_0 \leq \frac{c}{v}\,, \\ 1 & \text{if } n_0 > \frac{c}{v}\,, \end{cases} \tag{3.24}$$

where we can readily see that $n_1(c, n_0) \geq n_0$ if $n_0 \geq (c-a)/(v-a)$ and that this property holds for any time $t \geq 1$, i.e., $n_t(c, n_0) \geq n_{t-1}(c, n_0)$ if $n_{t-1} \geq (c-a)/(v-a)$.

Lemma 1. *For any time $t \geq 1$, the number of participants weakly increases if n_0 is sufficiently large, i.e.,*

$$n_t(c, n_0) \geq n_{t-1}(c, n_0) \quad \textit{if} \quad n_0 \geq \frac{c-a}{v-a} \ , \tag{3.25}$$

Considering the corner intervals in (3.24), Lemma 1 implies that if the initial participation satisfies $n_0 \geq c/v$ (if $n_0 < (c-a)/v$), the number of participants remains in the corner solution with $n_t = 1$ ($n_t = 0$) for any time $t \geq 1$.

We can now focus on the interior region where $(c-a)/v \leq n_0 \leq c/v$. To solve the first-order linear difference equation, $n_t = 1 - \hat{\theta}_t$, explicitly, we begin by expressing the number of participants at a protest at time $t = 1$ as

$$n_1(c, n_0) = 1 - \frac{c}{a} + n_0 \frac{v}{a} \ . \tag{3.26}$$

Next, consider the number of protest participants at time $t = 2$. By writing n_2 in terms of n_0 through the substitution of n_1, we obtain

$$\begin{aligned}
n_2(c, n_1) &= 1 - \frac{c - vn_1}{a} \\
\Rightarrow n_2(c, n_0) &= 1 - \frac{c}{a} + \frac{v}{a}\left(1 - \frac{c - vn_0}{a}\right) \\
\Rightarrow n_2(c, n_0) &= \left(1 - \frac{c}{a}\right)\left(1 + \frac{v}{a}\right) + n_0\left(\frac{v}{a}\right)^2 \ .
\end{aligned} \tag{3.27}$$

Analogously, the number of protest participants at time $t = 3$ is given by

$$\begin{aligned}
n_3(c, n_2) &= 1 - \frac{c - vn_2}{a} \\
\Rightarrow n_3(c, n_0) &= 1 - \frac{c}{a} + \frac{v}{a}\left[\left(1 - \frac{c}{a}\right)\left(1 + \frac{v}{a}\right) + n_0\left(\frac{v}{a}\right)^2\right] \\
\Rightarrow n_3(c, n_0) &= \left(1 - \frac{c}{a}\right)\left(1 + \frac{v}{a} + \left(\frac{v}{a}\right)^2\right) + n_0\left(\frac{v}{a}\right)^3
\end{aligned} \tag{3.28}$$

and more generally, for any subsequent time $t = k$, we have

$$\begin{aligned}
n_k(c, n_{k-1}) &= 1 - \frac{c - vn_{k-1}}{a} \\
\Rightarrow n_k(c, n_0) &= \left(1 - \frac{c}{a}\right)\left(1 + \frac{v}{a} + \left(\frac{v}{a}\right)^2 + \ldots + \left(\frac{v}{a}\right)^{k-1}\right) + n_0\left(\frac{v}{a}\right)^k \ .
\end{aligned} \tag{3.29}$$

Thus, the number of participants at time $t \geq 1$ can be written as

$$n_t(c, n_0) = n_0 \left(\frac{v}{a}\right)^t + \left(1 - \frac{c}{a}\right) \sum_{i=0}^{t-1} \left(\frac{v}{a}\right)^i \,, \tag{3.30}$$

and the change in the number of participants per period, for ease of expression, is denoted by $\Delta_t = \left(\frac{v}{a}\right)^{t-1} (n_1 - n_0)$.

(i) Under weak network effects ($v < a$), Equation 3.30 constitutes a geometric series in the limit as $t \to \infty$, converging to its supremum with $\lim_{t\to\infty} n_t(c, n_0) = (a - c)/(a - v)$.

(ii) Under strong network effects ($v > a$), the change in the number of participants Δ_t implies that in the limiting case in t, the participation level n_t converges to a corner solution. From Lemma 1, it must follow that $\lim_{t\to\infty} n_t(c, n_0) = 0$ if $n_0 < (c - a)/(v - a)$ and $\lim_{t\to\infty} n_t(c, n_0) = 1$ if $n_0 > (c - a)/(v - a)$, which finishes the proof. □

3.6.2 Variable Description and Robustness Appendix

Table 3.6: Variable Description ANES 2020.

Variable	Question
Protest Variable	
Protest	During the past 12 months, have you joined in a protest march, rally, or demonstration, or have you not done this in the past 12 months? *1. yes; 0. no.*
Socio-Demo. Characteristics	
Age	What is the month, day, and year of your birth? *Range 18-80. 80 indicating 80 or older.*
Education	What is the highest level of school you have completed? *1. Less than high school credential; 2. High school graduate; 3. Some college, no degree; 4. Associate degree - vocational; 5. Associate degree - academic; 6. Bachelor's degree; 7. Master's degree; 8. Professional school/Doctoral degree.*
Income	Please mark the answer that includes the income of all members of your family during the past 12 months before taxes. *Scale from 1. to 22. indicating income from 0-10k USD to >250k USD.*
Black	Are you Black? *1. yes; 0. no.*
Male	What is your sex? *1. Male; 0. Female.*
Health status	Self-evaluation of respondent's health. *0. poor; 1. fair; 2. good; 3. very good; 4. excellent.*
Married	Are you married? *1. yes; 0. no.*
Children	How many children age 0 to 17 live in this household? *0. None; 1. One; 2. Two; 3. Three; 4. Four or more*
Personal Attitudes	
Left-right-scale	Where would you place yourself on a scale from 0 to 10 where 0 means the left and 10 means the right?
Dev. Left-right-scale	Absolute number of 5 - previous answer
Religiosity	How important is religion in your life? *1. Extremely; 2. Very; 3. Moderately; 4. A little; 5. Not at all.*
Trust	Generally speaking, how often can you trust other people? *0. Never; 1. Some of the time; 2. About half the time; 3. Most of the time; 4. Always*
Longitude	Longitude of the respondent's state.
Latitude	Latitude of the respondent's state.
Political Attitudes Towards...	
Labor Unions	Feelings towards some of [these groups]. *Scale 0 to 100.*
Conservatives	Feelings towards some of [these groups]. *Scale 0 to 100.*
Liberals	Feelings towards some of [these groups]. *Scale 0 to 100.*
Feminists	Feelings towards some of [these groups]. *Scale 0 to 100.*
Christian Fundamentalists	Feelings towards some of [these groups]. *Scale 0 to 100.*
Police	Feelings towards some of [these groups]. *Scale 0 to 100.*
Black Lives Matter	Feelings towards some of [these groups]. *Scale 0 to 100.*

Note: Variable Description. Questions from the American National Election Survey 2020.

Table 3.7: Political Cluster Spatial Regression Results

Cluster	Right-Wing	Moderates	Left-Wing
Intercept	0.026	-0.022	0.072
	(0.046)	(0.113)	(0.057)
Male	0.000	-0.005	0.013
	(0.009)	(0.022)	(0.009)
Age	-0.001	0.000	-0.001
	(0.001)	(0.004)	(0.002)
Age^2	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Black	0.021	0.047	0.018
	(0.015)	(0.039)	(0.033)
Income	0.000	0.002	0.000
	(0.001)	(0.001)	(0.001)
Education	-0.002	0.006	0.000
	(0.002)	(0.004)	(0.002)
Left Right Scale	-0.003	-0.004	0.001
	(0.002)	(0.004)	(0.003)
Left Right Dev.	0.001	0.013**	0.005
	(0.002)	(0.005)	(0.004)
Health	0.010**	0.000	-0.005
	(0.005)	(0.011)	(0.005)
Married	-0.013	-0.042*	-0.017
	(0.011)	(0.025)	(0.011)
Number of children	-0.005	-0.017	-0.004
	(0.005)	(0.011)	(0.004)
Religious	0.001	-0.005	-0.003
	(0.002)	(0.005)	(0.003)
Trust government	0.002	0.011	0.001
	(0.003)	(0.009)	(0.005)
North East	0.001	-0.005	-0.005
	(0.010)	(0.022)	(0.014)
Mid West	-0.002	-0.023	-0.012
	(0.009)	(0.025)	(0.012)
South	-0.001	-0.018	-0.015
	(0.009)	(0.022)	(0.011)
Spatial Lag Y: λ	1.066***	0.979***	0.789***
	(0.168)	(0.131)	(0.297)
Spatial Lag ϵ: ρ	-0.884***	-0.866***	-0.556**
	(0.168)	(0.182)	(0.242)
Indirect effect/direct effect	*omitted*	39.91	3.53
Observations	2562	1368	2181

Heteroskedasticity robust spatial SE in parentheses; * p <0.1, ** p <0.05, *** p <0.01

Note: Political cluster full SARAR regression results based on Equation 3.14. Each column corresponds to a distinct political cluster.

Table 3.8: Robustness Baseline 2016

	Model (1) Neighbors: 10	Model (2) Neighbors: 15	Model (3) Neighbors: 20
Intercept	0.067*	0.056	0.052
	(0.040)	(0.034)	(0.032)
Male	-0.007	-0.007	-0.007
	(0.006)	(0.006)	(0.006)
Age	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)
Age^2	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Black	0.030**	0.029**	0.029**
	(0.014)	(0.014)	(0.014)
Income	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Education	0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)
Left Right Scale	-0.007**	-0.006**	-0.005**
	(0.003)	(0.003)	(0.003)
Left Right Dev.	0.010***	0.008**	0.008***
	(0.004)	(0.003)	(0.003)
Health	0.002	0.002	0.003
	(0.003)	(0.003)	(0.003)
Married	0.000	0.000	0.000
	(0.008)	(0.007)	(0.007)
Number of children	-0.005	-0.005	-0.004
	(0.004)	(0.003)	(0.003)
Religious	0.001	0.001	0.001
	(0.002)	(0.002)	(0.002)
Trust government	0.003	0.002	0.002
	(0.003)	(0.003)	(0.002)
North East	0.006	0.006	0.005
	(0.009)	(0.008)	(0.008)
Mid West	0.005	0.004	0.004
	(0.007)	(0.007)	(0.006)
South	-0.004	-0.004	-0.004
	(0.007)	(0.006)	(0.005)
Spatial Lag Y: λ	0.546**	0.723***	0.756***
	(0.271)	(0.238)	(0.240)
Spatial Lag ϵ: ρ	-0.335	-0.500**	-0.754***
	(0.226)	(0.227)	(0.247)
Observations	3266	3266	3266

Heteroskedasticity robust spatial SE in parentheses; * p <0.1, ** p <0.05, *** p <0.01

Note: Robustness for 2016 for the baseline SARAR regression results based on Equation 3.14. The full ANES sample is used, with a much smaller sample size for 2016 compared to 2020. Columns differ by the number of nearest neighbors used in the spatial weighting matrix (10, 15, and 20, respectively).

Chapter 4

From Silence to Salience – How Critical Events Shape Protests

Abstract: We propose a novel attention-based mechanism for protest participation, based on salience theory (Bordalo *et al.*, 2013). While existing models emphasize thresholds, networks, or informational dynamics, our model incorporates high-profile events that alter the context of protest decisions, shifting salience towards the benefits of protesting while underweighting associated costs. We challenge our model empirically by examining the killing of Black individuals by police officers as high-profile events and assessing their impact on attention and protest activity. Our findings show that such events significantly increase protest mobilization, with a substantial portion of the effect mediated by shifts in salience. This mechanism offers a behavioral component to existing protest models and expands our understanding of protest dynamics in the aftermath of attentional shocks.[1]

Keywords: Attention; Black Lives Matter; High-Profile Events; Protests; Salience Theory.

[1]This research article is a collaborative effort between Henrik Guhling and me, with each author contributing equally to the paper.

4.1 Introduction

High-profile events that attract substantial media coverage have the potential to trigger instant, far-reaching public uprisings in the form of protests. For instance, the murder of George Floyd on May 25, 2020, provoked large-scale protests against police violence and racial discrimination in the US and beyond. Like other protest movements during the last decade, protest mobilization went hand in hand with a wide online presence of the topic, steering societal attention towards the protest cause. These dynamics raise essential questions about how attention shocks triggered by high-profile events influence protest outcomes related to the affected issue. Exposure to a protest's cause through social media and other platforms alters the context in which individuals make participation decisions. By leading individuals to focus on the benefits of protesting and neglecting the associated personal costs, such exposure may increase mobilization–particularly in contexts marked by heightened public awareness following high-profile events, such as the murder of George Floyd.

Identifying the theoretical and empirical mechanisms through which high-profile events influence protest mobilization remains a significant challenge. From a theoretical perspective, while such events can change the context in which individuals decide whether to protest, modeling this context dependence is particularly challenging when neither the private benefits nor the private costs change in the event's immediate aftermath. From an econometric point of view, such events occur rarely, and thus examining them empirically poses challenges due to the typically small number of available observations. An increase in the number of included events almost certainly introduces either substantial heterogeneity among the considered events or across regional units of observation. Additionally, high-attention events are usually defined ex post, where the events have already undergone a selection process. This introduces a selection bias, as events are often included based on retrospective influence on protest outcomes rather than ex ante criteria.

To address the theoretical challenge, we develop an experience-based salience model that incorporates the attentional context into protest decisions, building on Bordalo *et al.* (2013, 2020a). Individuals decide whether to protest based on two attributes: the benefits (e.g., expressive utility, network effects) and the costs (e.g., opportunity costs, risk of violence). Over time, these individuals encounter experiences related to the protest or its underlying issue through channels such as social media, news outlets, or personal conversations. These experiences vary in how strongly they shape perceptions of the protest, with specific contexts amplifying their influence. For instance, extensive media coverage can make an incident significantly more impactful than a similar, less-publicized event. We refer to abrupt changes in this attentional environment as context shocks, which can shift attention toward the benefits of protesting and increase participation even without changes in the objective benefits or costs.

The proposed model connects high-profile events to protest mobilization by modeling an attention mechanism through salience. In doing so, it highlights the importance of contextual and attentional features in individual decision-making. We claim that the mechanism has substantial explanatory power for rapid protest mobilization patterns after high-profile events. Several instances clearly illustrate the mobilization dynamics of interest. Such events include, for example, the murder of George Floyd, which led to sustained anti-police violence protests; the Chinese extradition plans in 2019, which led to mass protests in Hong Kong; and the self-immolation of Mohamed Bouazizi in 2010, which started the Arab Spring.[2]

To overcome the aforementioned econometric difficulties, we shift the analysis to the local level in the US and define fatal police encounters with Black individuals as local context shocks. Especially in the context following George Floyd's death, such events received increased media attention and thus changed the context under which individual make their protest decisions. The investigation at the local level allows for assessing the effects of context shocks on the salience level and protest outcome as a direct comparison to units sufficiently distant but also sufficiently similar. Hence, we control for nationwide trends by comparing units with context shocks (treatment group) vis-à-vis units without context shocks (control group). We thereby circumvent the econometric difficulties of investigating nationwide context shocks by (a) defining events ex ante and (b) having suitable control units.

We test the model's predictions at the US county level for 2022; for some specifications, we extend the analysis period to 2017 through 2022. We apply treatment evaluation methods (regression adjustment, propensity score methods, and matching estimators) to obtain the effects of context shocks on salience levels and protest outcomes. Conditional independence is established using LASSO double-selection, and robustness for spillovers is tested using sample trimming. We presume Google Trends data to be a suitable proxy

[2]Several excellent studies investigate protest dynamics in these contexts: among others, Moreno-Medina *et al.* (2025) investigate biased media attention in police violence reporting, Williamson *et al.* (2018) assess the effect of police violence on protest outcomes in the context of BLM protests. Cantoni *et al.* (2019) argue that protests are strategic substitutes in the Hong Kong anti-authoritarian movement, and Cantoni *et al.* (2022) investigate which personality traits make individuals more likely to protest. In the Arab Spring literature, there are studies about its economic causes (Malik and Awadallah, 2013) and the role of social media (Khondker, 2011).

for the actual level of salience. To assess the dynamic predictions of the model, we further apply a distributed lag model.

The empirical evidence provides robust support for the major propositions derived from the model. Fatal police encounters involving Black individuals, interpreted as context shocks, are found to increase protest mobilization for Black Lives Matter (BLM) and anti-police violence demonstrations compared to a control group. The majority of this effect is attributable to the proposed salience mechanism. Furthermore, the mechanism is identifiable for the years 2018, 2020, and 2022, but differs substantially in size. After the George Floyd incident, the effects increased substantially, indicating that the nationwide context shock interacts with the local context shocks. In terms of sustained mobilization over the years, the salience mechanism does not seem to be the main driver. Instead, the salience mechanism appears to have substantial explanatory power for short-run mobilization processes, while the persistence of protest activity over time likely hinges on alternative mechanisms beyond attentional dynamics.

This paper makes several contributions to the literature. First, it introduces a novel attention-based channel that helps explain the emergence of protests. Second, it formalizes the role of high-profile events in decision-making by defining context shocks. Third, examining mobilization events at the local level allows for an assessment of their impact on protests by establishing an appropriate control group. This also makes it possible to classify such events by their magnitude and influence on protest activity. Finally, the paper offers an empirical approach to measuring context and salience by interpreting media coverage of a police shooting as context and Google Trends data as salience.

The paper's structure is as follows. After discussing the related literature in the remainder of this section, Section 4.2 presents the proposed model and testable hypotheses resulting from the model. In Section 4.3, we introduce our empirical strategy to assess the model's predictions. Section 4.4 presents the results and assesses their significance for the proposed model and the protest literature.

Related Literature The paper contributes to different strands of the literature. First, we add a novel attention mechanism to the literature that attempts to explain protests through economic theory, typically focusing on thresholds, networks, or information. The threshold model of Granovetter (1978) describes how an individual's decision to engage in collective action depends on the number or proportion of others already participating. Yin (1998) applies this framework to the domain of protests, showing that individuals may abstain from participation until a critical point is reached, after which joining becomes the rational choice. Such tipping points can potentially be reached by a small minority of "certain actors," which leads to widespread mobilization in the aftermath, as modeled by Wiedermann *et al.* (2020). While threshold models explain participation as a response to the number of others already protesting, our proposed attention mechanism focuses on why an individual decides to protest in the first place.

Another related strand of literature emphasizes the strategic interdependence of protest decisions. Shadmehr (2018) develops a model in which individuals decide whether to protest based on uncertain payoffs and the anticipated actions of others. They show that protest participation can exhibit either strategic substitutability or complementarity: when protest goals are modest, individuals act as strategic substitutes, while ambitious goals foster complementarity. Empirical evidence from Hong Kong supports the substitutability channel: Cantoni *et al.* (2019) find that students reduced their likelihood of participating in protests after learning that turnout would be higher than they initially expected. In contrast, in settings characterized by strategic complementarity, coordination becomes central. De Mesquita (2010) argues that public signals can shift beliefs about the broader anti-regime sentiment, thereby facilitating collective action. In contrast, our approach focuses on the decision of an individual independent of others' protest behavior, without relying on the protesting decisions of others, with heightened attention due to context shocks serving as the potential trigger.

Information-based models typically start with the assumption that individuals have imperfect information about government quality and/or the attitudes of fellow citizens. This category of models focuses on how information diffuses throughout society and how this mechanism affects protest mobilization. For example, Lohmann (1994, 2000) models protesting as a decentralized information aggregation mechanism and applies the theory to anti-regime mobilization dynamics in former East Germany. Little (2016) builds on that research by modeling how social media lowers the costs of communicating personal grievances with regimes and sharing information about protest coordination. Personal interactions are modeled in Barbera and Jackson (2019) in which individuals can update their beliefs according to communications with others. Belief updating about the government quality is explicitly modeled in Meirowitz and Tucker (2013) with a focus on the quality distribution of potential governments. Such informational changes are consistent with our proposed attention channel insofar as they alter the context in which individuals make decisions. Our analysis focuses on the attention mechanism, where context shocks influence individuals' perceptions.

Furthermore, this paper contributes to the empirical literature on protest occurrences, specifically to the literature on attention markets like social media platforms. Both González-Bailón *et al.* (2011) and Larson *et al.* (2019) identify a significant role of Twitter in the 2011 mobilization waves in Spain and the 2015 Charlie Hebdo protests in Paris. Enikolopov *et al.* (2020) demonstrate a similar effect of the VK platform, a major social media platform in Russia, on protests and their size. Regarding effect timing, Kann *et al.* (2023) use Twitter data and Black Lives Matter protests to show that social media increases interest in the topic in the short term, but that there is no sustainable increase in collective identity. In our analysis, we focus on the role of social media content in shaping the perceptual context in which individuals decide to protest, as captured by our attention mechanism framework.

Finally, we contribute to the strand of literature that extends the salience model of Bordalo *et al.* (2012, 2013) to topics in political economy. In this respect, Bordalo *et al.* (2020b) built on salience to explain why voters sometimes overweight extreme positions. In addition, Bilgin and Destan (2024) examines the impact of decoy candidates, who are not intended to be elected but can nonetheless influence the election outcome. Bonomi *et al.* (2021) show how the salience of a social divide leads voters to identify more strongly with their economic or cultural group, thereby pulling beliefs more strongly toward the stereotype of that group. Our analysis applies salience theory to the domain of protesting, using field data to measure both context and salience levels and to test hypotheses derived from the model.

4.2 Salience Model

To model our proposed attention mechanism, we adopt a salience-based framework. According to Bordalo *et al.* (2012, 2013), salience refers to the tendency of individuals to focus on the options, attributes, or outcomes that attract the greatest attention, i.e., the most salient ones. In salience theory, the context is typically defined by the choice set, with all available options forming a reference that shapes the perception of each option relative to the whole.

In contrast, our model focuses on a single choice option, namely protesting, in which perceptions are shaped by past experiences with similar protests or related topics. These past experiences include, among other aspects, information from social media, news outlets, personal conversations, and other sources. Thus, following Bordalo *et al.* (2020a), we model the context as an experience-dependent parameter, where the reference is formed not from all available options, but from the collection of past experiences with a single option, i.e., the protest.

In the following, we consider a decision maker (DM) who decides in each period $t = 0, 1, ...$ whether or not to take part in a protest $P = (b \geq 0, c \geq 0)$. Thereby, b reflects the potential benefits of the protest, and c the potential costs. The DM's outside option is given by $\underline{P} = (0, 0)$.

At time t, the DM retrieves all past experiences regarding the protest or the protest-related topic and aggregates them into a reference protest $\overline{P}_t = (\bar{b}_t, \bar{c}_t)$. Thus, the perception of the protest P at time t depends on $\overline{P}_t$. Each experience e_s, $s \leq t$, until time t is characterized by

$$e_s = (b_s, c_s, k_s) \in \mathbb{R}_0^+ \times \mathbb{R}_0^+ \times \mathbb{R}_0^+ \tag{4.1}$$

where b_s and c_s reflect the endogenously perceived benefits or costs of the protest at time s, respectively. Importantly, these perceived attributes need not coincide with the true attributes b and c. k_s denotes the exogenously given context at time s.[3]

Context is broadly defined to encompass the situational environment in which the DM perceives an experience. Such environments may include reading a newspaper, browsing social media, or attending a protest event. The way in which protest-related information is presented within these contexts can influence its perceived importance and relevance. For example, reading about a political development on the front page of a newspaper is likely to shape perceptions differently than encountering the same article in a less prominent section. Similarly, context can reflect the overall media coverage of a protest-related topic: frequent reporting can make the issue appear more important, while the choice of media outlet can frame the topic in a distinct way.

4.2.1 Similarity and Salience

All experiences until time t are organized in a memory database M_t, which is updated at each point in time. Then, the current experience at time t stimulates recall of similar past experiences.[4] Based on Bordalo *et al.* (2020a), similarity is measured along all three

[3] In general, we assume that different situations can be compared with each other. The context k_s captures this comparability by assigning different numbers to different situations. Initially, whether a value is high or low is irrelevant; however, the smaller the numerical distance between two values, the more similar the corresponding situations are assumed to be.

[4] Recall mechanisms are usually described by a cue that contains one or more dimensions of an experience and thereby evokes memories (Bordalo *et al.*, 2020a). However, in our case, a context at time t is always associated with perceived b_t and c_t, so that the cue always contains all three dimensions and simply corresponds to the current experience e_t at time t. Therefore, we omit the definition of a cue.

dimensions of experiences.

Definition 1 (Similarity)**.** *The similarity of the experience $e_s = (b_s, c_s, k_s)$, $s \leq t$, to the current experience $e_t = (b_t, c_t, k_t)$ is given by the multiplicatively separable distance:*

$$S(e_s, e_t) = S_1(|b_t - b_s|)S_2(|c_t - c_s|)S_3(|k_t - k_s|) \tag{4.2}$$

where $S_k : \mathbb{R}^0_+ \to \mathbb{R}_+$ is decreasing for $k = 1, 2, 3$.

Thus, $S(e_s, e_t)$ takes the highest value if the experience e_s contains the same parameter values as e_t. For example, $S(e_s, e_t)$ could have an exponential form, i.e.,

$$S(e_s, e_t) = exp(-\big[|b_t - b_s| + |c_t - c_s| + |k_t - k_s|\big]) \tag{4.3}$$

such that it takes values between 0 and 1.

Building on this concept of similarity, the reference protest is defined as the similarity-weighted average of perceived benefits and costs, where each experience in M_t is compared to the current one e_t, and experiences with greater similarity receive higher weights.[5]

Definition 2 (Memory weights and reference protest)**.** *The memory weight of experience $e_s = (b_s, c_s, k_s)$, $s \leq t$, at time t is given by*

$$w(e_s, e_t) = \frac{S(e_s, e_t)}{\sum_{e_u \in M_t} S(e_u, e_t)}. \tag{4.4}$$

Then, for the reference protest at time t, $\overline{P}_t = (\bar{b}_t, \bar{c}_t)$, it holds

$$\bar{b}_t = \sum_{e_u \in M_t} b_u \cdot w(e_u, e_t), \quad \bar{c}_t = \sum_{e_u \in M_t} c_u \cdot w(e_u, e_t). \tag{4.5}$$

In our salience model, the protest's attributes b and c are weighted based on discrepancies between their true values and the corresponding perceptions in the reference. Thus, at time t, the protest P is evaluated according to

$$U_t(b, c) = \frac{2 \cdot \delta^{-\sigma(b, \bar{b}_t)}}{\delta^{-\sigma(b, \bar{b}_t)} + \delta^{-\sigma(c, \bar{c}_t)}} \cdot b - \frac{2 \cdot \delta^{-\sigma(c, \bar{c}_t)}}{\delta^{-\sigma(b, \bar{b}_t)} + \delta^{-\sigma(c, \bar{c}_t)}} \cdot c \tag{4.6}$$

where $\delta \in (0, 1]$ measures the degree of salient thinking. For $\delta = 1$, there is no salience bias such that both attributes receive the same weight. The smaller δ, the stronger

[5]Unlike Bordalo *et al.* (2020a), we adopt a discrete-time framework because, in our model, sudden jumps in perceived benefits should have a substantial effect on the reference in order to capture context shocks. In continuous time and when the Riemann integral exists, such jumps could be measured as zero, as they may only represent single points. Using discrete time avoids this issue.

the perception bias. The salience function $\sigma(\cdot,\cdot)$ captures the key features of sensory perception (Bordalo *et al.*, 2013) summarized in Definition 3.[6]

Definition 3 (Salience function)**.** *For an attribute a, the salience function $\sigma(a,\overline{a})$ is a symmetric, continuous, and non-negative function that satisfies the following properties:*

1. ***Ordering**: Let $\mu = sign(a-\overline{a})$. Then, for any $\epsilon, \epsilon' \geq 0$ with $\epsilon + \epsilon' > 0$ it holds*

$$\sigma(a+\mu\epsilon, \overline{a}-\mu\epsilon') > \sigma(a,\overline{a}) \tag{4.7}$$

2. ***Diminishing sensitivity**: For any $a, \overline{a} \geq 0$ and all $\epsilon > 0$, it holds*

$$\sigma(a+\epsilon, \overline{a}+\epsilon) < \sigma(a,\overline{a}) \tag{4.8}$$

4.2.2 Context Shocks and Reference Protest

In our attention mechanism, we propose that high-profile events, such as the George Floyd incident, can alter the context substantially, acting as a context shock. This shock creates an experience with an extreme k_t and, consequently, an extreme b_t, which in turn shifts the reference protest. This shift changes salience, potentially directing more attention toward the protest's benefit b and prompting the decision-maker to participate in the protest.

To incorporate this mechanism in our model, we assume dependencies between the context k_t and the perceived benefit b_t. Furthermore, for tractability, we fix the cost perception bias.

Assumption 1. *We assume the following two conditions:*

(i) The perceived benefit at time t, b_t, depends on the context k_t, such that $b_t = f(k_t)$ with f continuous and injective.

(ii) The costs c and $\Delta_c = |c_t - c|$ are assumed to be constant over time, such that there is a constant cost perception bias.

The first assumption specifies a functional form for how the decision-maker's perception of benefits responds to changes in context, with larger changes in context leading to greater changes in perceived benefits.

[6] We adapt the rank-based salience approach of Bordalo *et al.* (2013) by introducing continuous attribute weights. This allows us to assess not only which attribute is more salient, but also the degree to which one exceeds the other in salience. Setting one attribute's salience to 0 and the other's to 1 recovers the original rank-based model.

The second assumption fixes the salience of costs, allowing all salience relationships between benefits and costs to be captured solely through changes in the salience level of b. This simplification involves no loss of generality, although perceived costs could, in principle, also be made context-dependent.

Next, we introduce a general definition of context shocks. Mathematically, a context shock occurs when the context reaches a level not previously achieved in a specific period of time.

Definition 4 (Context shock)**.** *A context at time t, k_t, is called a* ***context shock on*** *$\{s, ..., t\}$, $s < t$, if and only if the following property is fulfilled:*

$$\max_{u,w \in \{s,...,t-1\}} |k_u - k_w| < \max_{u,w \in \{s,...,t\}} |k_u - k_w| \tag{4.9}$$

If $s = 0$, then k_t is called a ***global context shock****.*

That is, we observe a context shock in t if there is a time period $\{s, ..., t\}$ such that the maximum fluctuation of the context becomes larger by adding the time t. In the simple case of constant context, every change represents a context shock. Importantly, it does not matter if k_t gets bigger or smaller. It is only vital that it changes.

Based on our mathematical definition, every context change can represent a context shock. However, this does not imply that every context shock also affects an individual's perception of protests. Only shocks that entail a change in the reference benefit $\bar{b}$ can change perceptions by influencing salience. The following lemma derives sufficient criteria for the relationship between context shocks and the reference benefit.[7]

Lemma 2 (Context shocks and reference benefit)**.** *Consider discrete time with $s \in \{0, ..., t\}$. Then a context shock at time t, k_t, induces a higher (lower) reference benefit, i.e., $\bar{b}_t > (<) \bar{b}_{t-1}$, if and only if the following inequality is fulfilled:*

$$\sum_{s=0}^{t} S(e_s, e_t) \cdot [b_s - \bar{b}_{t-1}] > (<)\ 0 \tag{4.10}$$

A global context shock k_t inducing $b_t = max_{s \leq t} b_s$ ($b_t = min_{s \leq t} b_s$) leads to a higher (lower) reference benefit if one of the following conditions is fulfilled:

(i) It holds $\bar{b}(k_{t-1}) \leq (\geq) \frac{1}{t} \sum_{s=0}^{t} b_s$.

(ii) For $\alpha_s = \frac{S(e_s,e_t)}{S(e_s,e_{t-1})}$, it holds $\alpha_i \geq \alpha_j \Leftrightarrow b_i \geq (\leq)\ b_j, \quad \forall i, j \in \{0, ..., t-1\}$.

[7] A proof of Lemma 2 can be found in Appendix 4.6.1.

Condition 4.10 is always fulfilled for a constant context before time t. Here, a global context shock always results in a higher (lower) reference benefit.[8] However, the reference benefit, and thus salience, remains unchanged if Condition 4.10 holds with equality. In the following, we define salience-changing context shocks as salience shifts.

Definition 5 (Salience shift)**.** *A context shock on* $\{s, ..., t\}$ *is called*

(i) ***positive salience shift*** *if and only if*

$$\sigma(b, \bar{b}_t) > \max_{u \in \{s,...,t-1\}} \sigma(b, \bar{b}_u) \tag{4.11}$$

(ii) ***negative salience shift*** *if and only if*

$$\sigma(b, \bar{b}_t) < \min_{u \in \{s,...,t-1\}} \sigma(b, \bar{b}_u) \tag{4.12}$$

A change in context that leads to the benefits' salience reaching a new level is referred to as a salience shift. When the advantages of protests are brought into greater focus, resulting in the benefit becoming more salient, this constitutes a positive salience shift. Conversely, if a context shock causes the salience of the benefits to diminish, thereby indirectly highlighting the disadvantages or costs of protests, this is termed a negative salience shift.

A salience shift can only occur through changes in the reference benefit of the form specified in Lemma 2. If the reference benefit already exceeds the true benefit b, the ordering property implies that even a slight upward change suffices to trigger a positive salience shift. If it lies below b, a slight downward change is sufficient. Theoretically, sufficiently large changes that move the reference benefit from above b to below it, or vice versa, can also induce a positive salience shift. The reverse logic holds for a negative salience shift.

Figure 4.1 illustrates how a context shock can influence the reference protest and, consequently, the perception of the benefit. With context and perceived benefit held constant, any change constitutes a context shock. According to Lemma 2, an upwards change of b_t in t_{shock} raises the reference benefit to a new, higher level. For $b_t = b$ with $t < t_{shock}$, this context shock heightens the benefits' salience, producing a positive salience shift. In the example shown, the context subsequently returns to its original level, causing b_t and thus the reference benefit to normalize to its prior level.

[8] Please note that the example function $S(e_s, e_t) = exp(-[|b_t - b_s| + |c_t - c_s| + |k_t - k_s|])$ from Equation 4.3 always fulfills Condition (ii) of Lemma 2 such that a higher b_t leads to an increased $\bar{b}_t$.

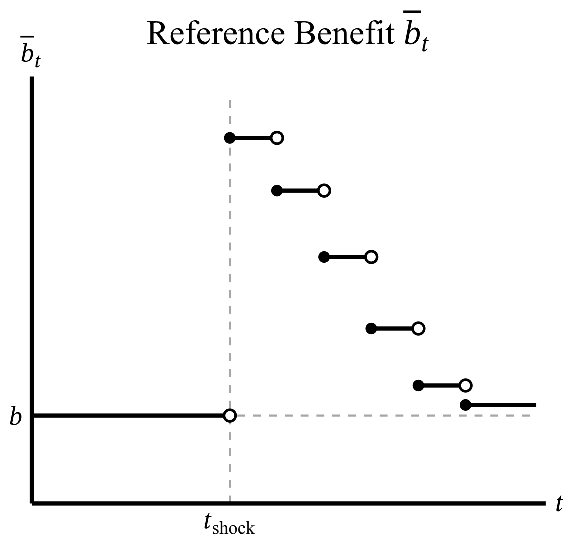


Figure 4.1: Shift in reference benefit $\overline{b}_t$ due to context shock in t_{shock} for $b_t = k_t$ and $S(e_s, e_t) = exp(-2|b_t - b_s|)$.

4.2.3 Protest Participation

Applying the model to the decision problem, a DM of type δ protests at time t if and only if

$$U_t(b, c) \geq U_t(0, 0) = 0. \tag{4.13}$$

Rearranging that leads to

$$\delta^{\sigma(c,\overline{c}_t)-\sigma(b,\overline{b}_t)} \geq \frac{c}{b}. \tag{4.14}$$

If both attributes are equally salient, Condition 4.14 shows that for $c > b$, nobody protests irrespective of type δ, whereas everybody protests for $b \geq c$. In this case, the protest decision only depends on the true attributes b and c.

In general, the participation of individuals in our model depends on which attribute is more salient. The more salient the benefit is compared to the costs, the more types of δ should participate. The following proposition gives an overview of protesting individuals based on the attributes' salience.

Proposition 3. *If the participation costs exceed the benefits, i.e., $c > b$, an individual of type $\delta \in (0, 1]$ decides to protest at time t if and only if $\delta \leq \delta_t^*$ where*

$$\delta_t^* = \begin{cases} 0 & \textit{if } \sigma(b, \overline{b}_t) \leq \sigma(c, \overline{c}_t), \\ \left(\frac{b}{c}\right)^{\frac{1}{\sigma(b,\overline{b}_t)-\sigma(c,\overline{c}_t)}} \in (0, 1) & \textit{if } \sigma(b, \overline{b}_t) > \sigma(c, \overline{c}_t). \end{cases} \tag{4.15}$$

If the benefits exceed the participation costs, i.e., $b \geq c$, an individual of type $\delta \in (0, 1]$

decides to protest at time t if and only if $\delta \geq \delta_t^*$ where

$$\delta_t^* = \begin{cases} 0 & \text{if } \sigma(b, \bar{b}_t) \geq \sigma(c, \bar{c}_t), \\ \left(\frac{c}{b}\right)^{\frac{1}{\sigma(c,\bar{c}_t)-\sigma(b,\bar{b}_t)}} \in (0,1) & \text{if } \sigma(b, \bar{b}_t) < \sigma(c, \bar{c}_t). \end{cases} \tag{4.16}$$

The proposition provides a threshold value δ_t^*, which depends on the attributes' salience at time t. Two trivial extreme cases arise by considering δ_t^*. First, if costs exceed benefits and are more salient, no one will protest. This makes sense, as attention is focused on the higher costs. Second, if benefits exceed costs and are more salient, all individuals will protest regardless of type δ. This also makes sense, since attention is on the higher benefits. In the remaining two cases, only a proportion of individuals protest, determined by their type δ.

If costs exceed benefits but the benefits are more salient, all individuals whose salient thinking exceeds a certain minimum, i.e., who focus sufficiently on the more salient attribute, will protest. As the salience of the benefit increases, the threshold value δ_t^* rises, leading to more types of δ participating.

Conversely, if benefits exceed costs but the costs are more salient, all individuals whose salient thinking is below a certain level, i.e., who do not focus too heavily on the more salient attribute, will protest. The smaller the salience gap between benefits and costs, the lower the threshold δ_t^* and the higher the participation rate.

Next, we examine how context shocks that induce a salience shift influence protest participation. The following corollary summarizes the results.

Corollary 1 (Context Shocks and Protest Participation)**.** *Consider discrete time with $s \in \{0, ..., t\}$.*

1. *For $c \geq b$, it holds:*

   (i) If $\sigma(b, \bar{b}_{t-1}) - \sigma(c, \bar{c}_{t-1}) \geq 0$, then every ***positive salience shift*** *on $\{s, ..., t\}$ leads to increased protest participation.*

   (ii) If $\sigma(b, \bar{b}_{t-1}) - \sigma(c, \bar{c}_{t-1}) < 0$, then only sufficiently large ***positive salience shifts*** *on $\{s, ..., t\}$ lead to increased protest participation.*

   (iii) If $\sigma(b, \bar{b}_{t-1}) - \sigma(c, \bar{c}_{t-1}) > 0$, then every ***negative salience shift*** *on $\{s, ..., t\}$ leads to decreased protest participation.*

2. *For $c \leq b$, it holds:*

*(iv) If $\sigma(b, \bar{b}_{t-1}) - \sigma(c, \bar{c}_{t-1}) \leq 0$, then every **negative salience shift** on $\{s, ..., t\}$ leads to decreased protest participation.*

*(v) If $\sigma(b, \bar{b}_{t-1}) - \sigma(c, \bar{c}_{t-1}) > 0$, then only sufficiently large **negative salience shifts** on $\{s, ..., t\}$ lead to decreased protest participation.*

*(vi) If $\sigma(b, \bar{b}_{t-1}) - \sigma(c, \bar{c}_{t-1}) < 0$, then every **positive salience shift** on $\{s, ..., t\}$ leads to increased protest participation.*

The results indicate that not all context shocks that alter salience necessarily affect protest participation. When costs exceed benefits and are also more salient, only a substantial increase in the salience of benefits can spur participation. The George Floyd incident exemplifies such a powerful shock.

Conversely, when benefits exceed costs and are more salient, a sharp decline in benefits' salience is needed for costs to become salient again. Severe repressions in autocratic regimes, such as the Tiananmen massacre in China, illustrate such large negative salience shifts. However, multiple smaller negative shifts can also erode participation over time. For instance, waning attention to the protest-related topic can cause movements to shrink and eventually vanish, unless a new context shock triggers a positive salience shift that recaptures attention.

Discussion of Protest Movements over Time Finally, we turn to how our model can explain the life cycle of protest movements. Before protests emerge, mobilization is typically low, suggesting that costs exceed benefits ($c > b$) and are more salient. In our framework, this can also occur when $b > c$ but the salience of benefits is sufficiently low, drawing attention to costs. In both situations, positive salience shifts can increase participation, though in the first case, the shift must be sufficiently large. Major salience shifts, such as the George Floyd incident, often serve as trigger events for protest movements. In some cases, several smaller shifts can cumulatively spark mobilization, albeit more gradually. In many cases, however, a single major event can be identified as the catalyst.[9] After such a shock, additional smaller shocks can further boost participation.

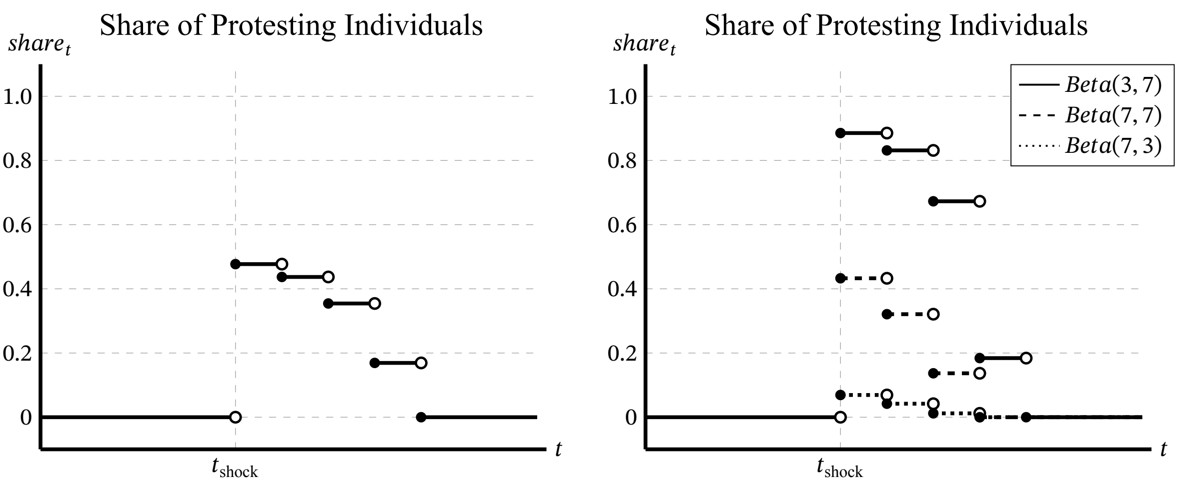
Over time, attention typically dissipates or shifts to other issues, reducing the salience level. As salience falls, participation declines, eventually returning to its initial low level, leaving only those individuals who were consistently motivated. Strong negative salience

[9]Examples include the George Floyd incident; the self-immolation of Mohamed Bouazizi, which heralded the Arab Spring; the shooting of Neda Agha-Soltan, which ignited the Green Movement in Iran; and the death of Mahsa Amini, which triggered the Women's Rights Movement in Iran.

shifts, such as repression, can accelerate this decline and bring the protest "wave" to an earlier end.

In principle, the ratio of b to c can change from $c > b$ to $b > c$ and back, for example, through increasing network effects and a growing protest movement. Yet even in this case, a decline in the salience of benefits reduces participation, leaving only a small core of participants. However, the case starting with $c > b$ is particularly interesting: if benefits are more salient and the ratio changes to $b > c$, a subsequent drop in benefits' salience means that the movement's initiators, i.e., those with extremely salient thinking, are also the first to withdraw. This pattern is observable in movements that radicalize over time, as initiators disengage when the protest shifts in a direction they no longer support.

Figure 4.2 presents a numerical example of protest participation for different distributions of δ. In this example, a positive salience shift produces an immediate jump in participation, but over time, negative salience effects dominate, leading to a decline. For $\delta \sim \mathcal{U}(0, 1]$, the participation share matches the critical type δ_t^*. Under $\delta \sim Beta(\alpha, \beta)$, there is less probability mass on the tails of the distribution, accelerating the decline. In the left-skewed case ($Beta(3, 7)$), the initial surge in participation following the context shock is much larger, while in the right-skewed case ($Beta(7, 3)$), the effect is smaller. With a symmetric $Beta(7, 7)$ distribution, the initial increase is similar to the uniform case, but the subsequent decline is also faster.[10]



(a) $\delta \sim \mathcal{U}(0, 1]$

(b) $\delta \sim Beta(\alpha, \beta)$

Figure 4.2: Share of protesting individuals for $b = 0.7$, $c = 1$, and $\sigma(b, \bar{b}) = \frac{|b-\bar{b}|}{|b|+|\bar{b}|}$.

[10]The corresponding $Beta$ densities are shown in Appendix 4.6.2, Figure 4.6.

4.2.4 Hypotheses

Based on our theoretical model, we derive empirically testable hypotheses. For the empirical analysis, we treat fatal police encounters with Black individuals as context shocks and focus on mobilization patterns in BLM and anti-police protests between 2018 and 2022. For simplicity, we refer to the salience of the protest's benefit as the salience of the protest, as in our theoretical model, variation in protest behavior arises solely from changes in the salience of the benefit, with higher salience directly increasing the likelihood of participation. Furthermore, due to limited data quality in terms of the number of protest participants, we rely on the number of protest events at the US county level as a dependent variable.[11]

Our first hypothesis reflects the core idea that certain acts or events can serve as context shocks, thereby influencing protest outcomes.

Hypothesis 1 (Context shock effect). *We expect context shocks to impact protest outcomes, i.e., police shootings significantly increase the number of anti-police and Black Lives Matter protest events.*

The second hypothesis aims to test the role of salience in shaping protests. In our model, a context shock constitutes a positive salience shift if it raises salience, which in turn increases the number of protest events.

Hypothesis 2 (Salience mechanism). *We expect that context shocks impact salience, i.e., the salience of protest movements is shifted significantly upwards by police shootings (Mechanism 1). Furthermore, this increased salience leads to more protest events, i.e., the number of protest events increases significantly by a shifted salience (Mechanism 2).*

The third hypothesis tests how differently sized context shocks influence protests. If larger shocks lead to more salience, they should also produce a stronger increase in protest events. Since we classify shocks by the number of Google search results for a given shooting, greater media coverage should correspond to higher salience and, consequently, more protests.

Hypothesis 3 (Context shock size). *If bigger context shocks lead to bigger positive salience shifts, we expect to observe more protest events when the context is shocked more drastically.*

[11] Although the used database (Crowd Counting Consortium, 2021) reports estimates of protest participation, these values vary widely between low and high estimates and are missing for a substantial share of events. Therefore, we rely on the number of protest events as a more consistent and informative measure for protest mobilization.

The final hypothesis examines the dynamic implications of our model. If the context returns to its pre-shock level, the impact of a context shock should fade as salience also normalizes.

Hypothesis 4 (Dynamic effects). *If context normalizes over time, we expect to observe a decreasing effect of context shocks on salience and the number of protest events.*

4.3 Empirical Strategy

The following section describes the data collection procedure and empirical strategy to test the theory-derived hypotheses. In the first step, the most important variables for context shocks, salience, and protest outcomes are collected and introduced. Then, three empirical estimation strategies are presented: (1) binary treatment evaluation, (2) continuous treatment evaluation, and (3) distributed lag model (DLM). These estimation strategies are applied in different parts of the results section to test the hypotheses. The final subsection addresses challenges in the empirical strategy and presents sensitivity tests to demonstrate the robustness of the estimated effects.

4.3.1 Measuring Context Shocks, Salience, and Protests

The three key variables to empirically test the model's predictions are fatal police encounters as context shocks, salience, and protest outcomes. The question of interest is whether there is evidence in the data that a context shock impacts the salience of protesting, which in turn is hypothesized to impact protest mobilization. As protest coordination necessarily involves social groups, the US county level is selected as the observational unit. For the cross-sectional analysis, we choose 2022 as the base year to maintain some distance from the atypical year 2020, marked by nationwide protests following George Floyd's murder. Throughout the paper, we also apply estimation techniques to exploit the time variation in the variables in the period from 2018 to 2022.

Context shocks are measured as fatal encounters of Black individuals with the police (from here on referred to as *fatal police encounters*). The *Mapping Police Violence* database provides detailed data on all fatal encounters of the police with civilians. The data are filtered for fatal encounters with black civilians only and subsequently aggregated on the US county level.[12] These events constitute (micro-level) context shocks, which are

[12] In 2022, there were 53 counties in which more than one fatal police encounter happened. In the first step, we do not account for this heterogeneity and classify all counties with at least one fatal police

supposed to facilitate protest coordination. Having the database in place, the racial disparities in exposure to police violence become visible. While Black individuals' population share is 12.6%, their share in fatal encounters with the police is 23.5%. The states with the highest per capita police killings are North Carolina (0.25 per 1,000 residents), New York (0.23 per 1,000 residents), Nevada (0.20 per 1,000 residents), and New Mexico (0.10 per 1,000 residents).

To capture the extent to which a fatal police encounter led to a context shock, we rely on the Google SerpAPI. Based on this, we determine for each police shooting the number of Google results that contain the exact name of the victim and at least one of the keywords "police shooting", "police killing," "fatal police encounter," or "officer-involved shooting." In addition, only results that were published after the date of the victim's death are taken into account. This measure of media coverage about the fatal police encounters serves two purposes. First, it is a more precise measure of how much the context was shocked in a given county. Second, it aggregates the number of results when there was more than one fatal police encounter in a county in a given year (for 2022, for example, there were 53 counties in which more than one fatal police encounter happened).

We measure the attentional response to context shocks as salience, proxied by Google Trends data. These data reflect the relative frequency of searches for the term "police shooting." The intuition behind this measure is that the more often the keywords "police shooting" are googled, the more attention people pay to police brutality and problematic race-based discrimination. Importantly, we measure the extent of a context shock by the media coverage of the event, whereas salience is captured by individuals' active search behavior, which we interpret as an indicator of attention. For a discussion on measuring issue salience with Google Trends data, see Mellon (2013). In other scientific contexts, Google Trends data is frequently used to obtain measures of economic uncertainty (Castelnuovo and Tran, 2017; Choi and Varian, 2012). Using Google Trends, we obtain the relative frequency of how often the term was googled for a yearly cross-section on a regional level. The region with the highest Google searches receives a value of 100, and all other values are relative to this benchmark. The regional units of the Google Trends dataset contain multiple counties, and their values are assigned to the lower-level county units. For robustness, all estimations are also conducted on the higher-level Google Trends regional units. The results are generally robust to either specification.

Obtaining a valid measure for salience over time is more elaborate because the Google

encounter as treated counties do. Later, we specify media coverage and aggregate a new variable at the county level, which accounts for heterogeneities between treatment intensities.

API sets the maximum value of each region automatically to 100 and then displays all other values relative to the maximum. Cross-county comparisons are, therefore, infeasible because each county, at some point, has a maximum value of 100. We follow Stephens-Davidowitz (2014) to overcome this problem by using reference words that are as uncorrelated to time and geography as possible and measure the salience relative to these reference words. Importantly, the reference words set the maximum value, and the relative frequency of the salience measure is thus comparable over time and across all counties. The selected reference words include “recipies” (misspelled on purpose), “Restroom”, “German”, and “Back Pain”. Figure 4.3 illustrates how salient the term “Police shooting” was throughout the period of investigation.

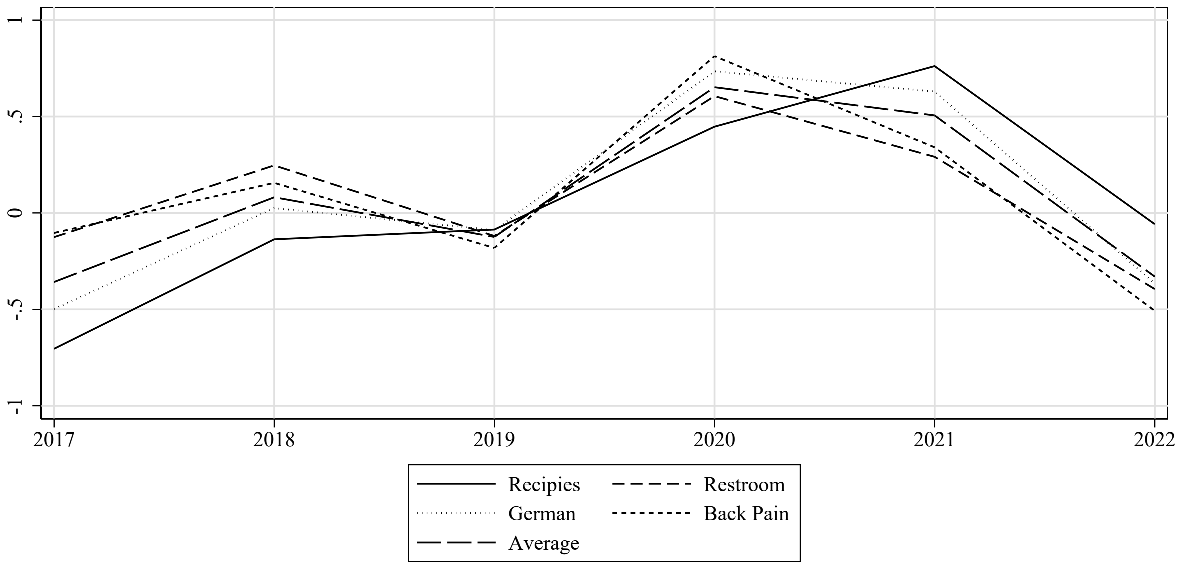


Figure 4.3: Salience of “Police Shooting” over time across all counties. Note that “Recipies” is misspelled on purpose.

Protest data are retrieved from the *Crowd Counting Consortium* database (Crowd Counting Consortium, 2021) – a dataset based on news data. Within the covered years from 2017 to 2022, a natural language processing algorithm retrieves information about protest events from a large number of media outlets. The result is an event database with geocoded protest events across the US. The protest events are aggregated on the US county level and filtered for protest events whose motivations are classified as “race relations” or “against police.” These two motivations coincide with a large share of protest events, and as soon as one category applies, the event is included in the aggregation.

Table 4.1 describes the main variables of interest. There was a fatal police encounter in 5.9% of US counties in 2022. The average salience of “Police Shooting” was 40. This is measured relative to the maximum of 100 in the county with the highest salience level. Furthermore, in 11% of the counties, there was at least one protest event about race relations or police violence. The remaining empirical section describes techniques by which we aim to establish a causal link from the context shock (fatal police encounter) to protest events via salience.

Key Variables	N	Mean	Std. Dev.	Min	Pctl. 25	Pctl. 75	Max
Fatal Police Encounters	3,222	0.059	0.24	0	0	0	1
Salience "Police Shooting"	3,045	40	9.9	18	34	44	100
BLM Protest Events	3,222	1.4	18	0	0	0	573
BLM Protest Events (binary)	3,222	0.11	0.31	0	0	0	1

Table 4.1: Summary Statistics for 2022. "BLM Protest Events" include all protest events to which the category "against police" or "race relations" applies. As explained in the text, fatal police encounters include only incidents with black victims.

4.3.2 Estimation Strategy

We employ a set of cross-sectional estimators to test the implications of Hypotheses 1-3. As a first approach, we assess the effect of context shocks on the number of BLM and anti-police protests on the US county level (Hypothesis 1). Then, we test the mediation channel via salience by splitting the channel into two mechanisms, assessing first the effect of context shocks on salience, and then of salience on protest outcomes (Hypothesis 2). The third specification evaluates heterogeneities towards the size of context shocks by relying on the number of Google results (Hypothesis 3). Finally, the dynamic effects of the salience mechanism are assessed (Hypothesis 4).

I. Binary Treatment

The target parameter for estimation is the average treatment effect on the treated (ATET). This parameter gives an indication of how much a context shock has increased the salience or protest outcome in counties in which a fatal police encounter occurred. The average treatment effect is of less interest as there are some counties with certain characteristics (e.g., almost exclusively white population) that make fatal police encounters (of Black individuals) so unlikely that it is questionable to extrapolate from the treated observations to these counties.[13] Equation 4.17 presents the target parameter where Y denotes the dependent variable, X denotes the covariates, and W denotes the binary treatment variable.

[13]This fact becomes visible in Figure 4.7. There, a large number of untreated counties have a treatment probability of close to zero.

$$\tau_{\text{ATET}} = E[E[Y \mid X, W = 1] - E[Y \mid X, W = 0] \mid W = 1] \tag{4.17}$$

The identifying assumptions underlying the estimators are conditional unconfoundedness and common support (Abadie and Cattaneo, 2018). To alleviate potential concerns for a violation of conditional unconfoundedness, the LASSO-selection procedure is applied as suggested by, for example, Belloni *et al.* (2014). From a set of 76 potential control variables, only those that have predictive power for either the treatment or outcome enter as control variables. All relevant details about the LASSO selection procedure are presented in the Appendix.[14]

Besides the technical assumption that the treatment probability cannot be 1 (Wooldridge, 2010), we assess the overlap of the treatment and control groups to ensure that there are valid units for comparison in the opposite group. The propensity score is estimated as a logit regression of the treatment variable on a set of covariates. The propensity score density function is displayed in Figure 4.7 in Appendix 4.6.4. Much of the density of the control group is centered at 0, with very low propensity scores. As the target parameter is the ATET, there is no reason to worry because we want to assess whether there are comparison units for the treatment group. These comparison units are given as the control group shows substantial density for medium to high values of the propensity score. In case there remain concerns about extreme propensity scores, we follow Crump *et al.* (2009) in trimming the sample by excluding all observations with propensity scores lower than 0.05 or higher than 0.95 in the sensitivity analysis.

There are various estimators to obtain our target parameter τ_{ATET} as presented in Abadie and Cattaneo (2018). Hence, we apply five estimators to demonstrate our results' robustness: (1) regression adjustment, (2) inverse probability-weighting, (3) inverse probability-weighting regression adjustment (doubly-robust estimation), (4) covariate matching, and (5) propensity score matching. All of these estimators ensure that under the conditional independence assumption, the target parameter is statistically identified. This objective is either achieved by weighting or matching on propensity scores obtained by a prior logit regression or by directly adjusting the comparison unit using a regression adjustment approach. Additionally, we employ a covariate matching estimator that compares units according to how similar they are based on the selected covariates. The included covariates are selected using the previously described LASSO selection procedure.

[14] Appendix 4.6.3 offers an extensive description of the applied procedure. Appendix 4.6.3 presents an exhaustive list of all 76 potential control variables. Finally, Appendix 4.6.3 lists the sets of covariates used, which will be referenced in all subsequent regression tables.

A detailed account of the estimators is delegated to Appendix 4.6.4.

II. Continuous Treatment

Investigating the effect of salience on protest outcomes requires adapting the estimation strategy for continuous treatments (Mechanism 2 of Hypothesis 2). For this, we rely on Imbens (2000) and Hirano and Imbens (2004). The authors developed a method based on a generalized propensity score (GPS) that accommodates a continuous treatment variable.

The generalized conditional unconfoundedness assumption implies that the potential outcome for all possible treatment dosages must be independent of the actual received treatment dosage, conditional on a set of control variables. Imbens (2000) demonstrates that, similar to the propensity score method discussed in the previous subsection, conditioning on the GPS suffices to eliminate all bias arising from the non-random treatment dosage assignment. The GPS is a scalar function that indicates the likelihood of receiving the actual treatment dosage based on a set of covariates.

The estimation procedure follows Hirano and Imbens (2004) and their application of the GPS for assessing continuous treatments. In the first step, we regress the continuous treatment on a constant and a set of control variables using OLS. From this model, we obtain predicted values for each observation and the standard deviation of the residuals. The GPS for each observation is derived by assuming a normal distribution of the disturbances. Therefore, we use the conditional Gaussian density to estimate the GPS ($\hat{r}(w, X_i)$) for each observation.

With the GPS at our disposal, we follow Hirano and Imbens (2004) in modeling the treatment effect using a quadratic conditional mean model. If conditional unconfoundedness holds, the model fully accounts for all potential confounding variations in the covariate set. It includes a quadratic term and an interaction term between the treatment dosage and the GPS to impose weaker assumptions about the functional form of the treatment effect. The model is estimated using OLS regression, as illustrated in Equation 4.18. Due to the uncertainty in the first-stage estimation of the GPS ($\hat{r}(w, X_i)$), standard errors are obtained using the Delta method.

$$Y_i = \alpha_0 + \alpha_1 w + \alpha_2 w^2 + \alpha_3 \hat{r}(w, X_i) + \alpha_4 \hat{r}(w, X_i)^2 + \alpha_5 w \hat{r}(w, X_i) + \epsilon_i \tag{4.18}$$

Once all estimations are conducted, we obtain a dose-response function and plot it, including its confidence bands. Based on this function, we can assess the treatment effect of salience on protest outcomes at each level of the salience distribution.

III. Distributed Lag Model

To assess the dynamic persistence of context shocks, we estimate a distributed lag model (DLM) presented in Equation 4.19. This contributes to assessing the intertemporal dynamics of context shocks as theorized in Hypothesis 4. We are interested in the lagged effects of context shocks on (1) protest mobilization and (2) salience level. Therefore, W is a binary variable indicating whether there was a fatal police encounter in a given year or not. The outcome variable Y is the (transformed) protest variable, and, in the second specification, the salience level. Equation 4.19 is then estimated using an OLS regression with heteroskedasticity-robust standard errors. Since salience is observed at a higher regional level, the results for the effect of context shocks on the current salience level are displayed for both the county level and the Google Trends regional level.

$$Y_i = \alpha_0 + \alpha_1 W_{y,i} + \alpha_2 W_{y-1,i} + \alpha_3 W_{y-1,i} + ...\beta X_i + \epsilon_i \tag{4.19}$$

The overall effect of context shocks on both the protest outcomes and salience level is assessable by computing the long-run propensity (LRP). The LRP is the sum of all lagged coefficients for which we are able to compute standard errors, too. These are obtained in a regression of the outcome variable on a constant, the contemporaneous treatment, the first difference of all past treatments, and a set of controls. The contemporaneous treatment in this regression represents the LRP, including a valid standard error to assess the joint significance of all treatments (Wooldridge, 2016).

4.3.3 Robustness and Identification Considerations

To complement the main estimation results, we conduct a series of robustness checks, with detailed results presented in the Appendix and referenced throughout the results section. First, we evaluate covariate balance and placebo treatment effects to assess the credibility of the identification strategy. Second, we outline our approach to account for potential spatial spillovers, which may bias treatment effects if left unaddressed. To address empirical concerns, we implement alternative specifications designed to mitigate two key issues: the influence of outliers and heterogeneity across margins of the dependent variable. The strategies employed to address these concerns are discussed in this subsection. We omit further discussion of confounding and overlap, as these issues have already been addressed through LASSO-based covariate selection and sample trimming procedures described earlier.

(1) Covariate Balance. Table 4.8 in the Appendix 4.6.4 presents balance diagnostics

and placebo tests for the baseline specification. Standardized mean differences (SMDs) are used to assess the extent to which each estimator achieves covariate balance between the treatment and control groups. Overall, most covariates exhibit satisfactory balance across estimators. However, notable discrepancies remain for the *Gini coefficient* and *collective efficacy*, which show residual differences in means between groups. To formally assess balance, we employ the overidentification test proposed by Imai and Ratkovic (2014), which fails to reject the null hypothesis of covariate balance, thereby supporting the comparability of treatment and control groups under the preferred specifications.[15] In addition, results from the placebo treatment test suggest that the IPW, IPW-RA, and propensity score matching estimators outperform the RA and covariate matching estimators in mitigating confounding bias. Consequently, estimates derived from the RA and covariate matching approaches should be interpreted with caution, while those from the weighting and propensity score matching estimators appear more credible and robust.[16]

(2) Spatial Spillover. There are two kinds of spatial spillovers that affect protest mobilization and salience levels across counties: nationwide and regional spillovers. The first category of nationwide spillovers does not pose a challenge to the empirical estimation strategy. The most striking event that caused spillovers across the nation and beyond has been the murder of George Floyd by a police officer. Such nationwide spillovers are interesting as they shift the context but do not pose a problem for effect identification, as both the control and the treatment groups are affected by them to the same extent. Hence, the level of protest events due to the nationwide context shock increases similarly in the treated and untreated counties and, thereby, cancels out in the estimation of the ATET.

Regional spillovers present a challenge for empirical estimation because a fatal police encounter in one county can influence neighboring counties through social connections, regional media, or social media, undermining their validity as a control group. This violates the Stable Unit Treatment Value Assumption, as the outcomes in one county may be affected by events in nearby areas. To address this concern, the study follows a sample trimming approach from Clarke (2017), excluding counties within a certain distance of treated counties to form a "clean" control group unlikely to be impacted by spillovers. For robustness, additional regressions exclude counties within a 75- and 100-mile radius to

[15] This test is only applicable for weighting estimators. The corresponding p-values of the overidentification test for IPW and IPW-RA are 0.80 and 0.89, respectively.

[16] We attribute the residual significance of the placebo treatment for the IPW, IPW-RA, and propensity score matching estimators in 2021 to the time persistence of protest events across years. Specifically, the correlation of protest events in 2022 with those in 2021, 2020, and 2019 is 0.88, 0.69, and 0.62, respectively.

test the result's sensitivity to spatial spillovers.

(3) Outliers. A source of empirical imprecision, not necessarily bias, is extreme values in the outcome variable. Possibly, the observed effects could only hold for a few treated counties that coincide with the outliers of protest mobilization. To assess whether the described mechanisms hold in general, we conduct robustness checks for a transformed dependent variable. Therefore, all estimations with the number of protest events as a dependent variable are also estimated with the inverse hyperbolic sine transformation (IHS, $arsinh\,(x) = ln(x + \sqrt{x^2 + 1})$) of the dependent variable. This transformation is an approximation of the log transformation, but is also valid for zero values. Furthermore, the transformation allows for an (approximate) interpretation of the results in percentages.

(4) Extensive Margin. Another strategy to reduce the influence of outliers and assess the extensive margin is to transform the dependent variable into a binary variable. The binary variable indicates one if there was at least one protest event in the given county and zero otherwise. The estimation, thus, assesses the change in the likelihood that at least one protest occurs as a consequence of treatment. This helps us to evaluate whether the results are only driven by the intensive margin or whether the treatment also has an impact on the extensive margin.

4.4 Empirical Results

The results are presented in the following manner, progressing from the most general to the most specific hypothesized relationships. We begin by estimating the overall impact of context shocks on protest outcomes using the baseline static model (Hypothesis 1). Next, we examine whether salience serves as a mediating mechanism between context shocks and protest activity, thereby testing Hypothesis 2. Building on these findings, we evaluate Hypothesis 3, which posits that larger context shocks exert stronger effects on both mobilization and salience of protesting. Time dependencies are assessed relying on the DLM estimation strategy (Hypothesis 4). A comprehensive set of robustness checks in Subsection 4.4.5 addresses potential concerns related to model specification and identification.

4.4.1 Context Shocks and Protest Outcomes

We begin by discussing the evidence collected to answer Hypothesis 1, which theorizes that a context shock should impact protest mobilization. As outlined in Subsection 4.3.2, we investigate whether a context shock caused BLM or anti-police protest events. Additionally, a sensitivity analysis is performed in the appendix to demonstrate robustness.

The results presented in Table 4.2 indicate that there is a substantial effect of a context shock on protest mobilization. For those counties that are treated, the number of protest events increased by 8.5-10.5 units as a consequence of a context shock. Given that the mean of BLM and anti-police protest events in all counties for 2022 is 1.4, this is a substantial increase and highlights the importance of the surrounding context for protest coordination.

Table 4.2: Binary Treatment Evaluation: Context Shock on Protest Outcomes 2022

	RA	IPW	IPW RA	C. Match	P.S. Match
Context Shock	10.463**	8.768*	8.511*	9.044**	9.728**
	(4.364)	(4.474)	(4.456)	(4.525)	(4.292)
Observations	3007	3007	3007	3007	3007

Heteroskedasticity robust SE in parentheses; * p <0.1, ** p <0.05, *** p <0.01

Note: Estimation results of the Average Treatment Effect on the Treated (ATET) from the context shock on protest events for the year 2022. LASSO selection covariate Set 1 is applied.

The results demonstrate considerable robustness, as shown by the sensitivity analyses reported in Table 4.9 in Appendix 4.6.5. Estimates using the transformed dependent variable suggest that a context shock increases the number of protest events by 40-73.4 percent and raises the probability of at least one protest occurring in a county by 10.5-24.2 percentage points (all effects are estimated on the treated units). These findings remain robust to sample trimming procedures designed to address potential spillover effects and violations of the overlap assumption. The larger coefficient in the spatial spillover specification indicates the presence of spillover effects. In particular, counties adjacent to treated units appear to experience more protests as a consequence of their neighboring counties being treated, which may inflate protest levels in the control group. These spatial spillovers likely bias the baseline estimates downward by attenuating the difference between treated and control units.

The results, along with the robustness checks, support Hypothesis 1, providing confidence that fatal police encounters qualify as context shocks for BLM and anti-police protests. The hypothesis states that context shocks generally impact protest mobilization. We observe significant results across several specifications and robustness tests, reinforcing Hypothesis 1.

4.4.2 Salience Mechanism

After establishing the initial link between a context shock and protest mobilization in the previous subsection, we focus on the explicit salience mechanism in the following paragraphs. As modeled in Section 4.2, context shocks are theorized to affect the protest outcome by changing the salience of protesting. Hypothesis 2 is thus divided into two mechanisms. Mechanism 1 implies that a fatal police encounter increases the salience in a given location. The subsequent Mechanism 2 hypothesizes that increased salience causes more protesting. Both mechanisms are empirically assessed by relying on binary and continuous treatment evaluation techniques presented in Subsection 4.3.2.

Mechanism 1 is robustly identified across a variety of estimation techniques. The standard binary treatment technique in the first line of Table 4.3 indicates that a context shock increases the salience level by 2.6-3.5 units. This corresponds to a salience increase of 26-35% of the standard deviation of the salience measure.[17] The effect of context shocks on salience remains robust to the transformation of the dependent variable, accounting for spillovers, addressing overlap, and shifting the observational unit to the Google Trends regions (Table 4.10 in Appendix 4.6.5).

Table 4.3: Static Treatment Evaluation: Context Shock on Salience

	RA	IPW	IPW RA	C. Match	P.S. Match
Context Shock	2.612***	3.388***	3.464***	2.704***	2.979***
	(0.830)	(0.868)	(0.909)	(0.858)	(0.757)
Observations	2924	2924	2924	2924	2924

Heteroskedasticity robust SE in parentheses; * p <0.1, ** p <0.05, *** p <0.01

Note: Estimation results of the Average Treatment Effect on the Treated (ATET) from the context shock on salience for the year 2022. LASSO selection covariate Set 1 is applied.

Mechanism 2 is also identified relying on the estimation technique for continuous treatments presented in Subsection 4.3.2. Figure 4.8 shows the dose-response function of each salience level for its effect on protest mobilization. At the mean level (salience equals 38), a one-unit increase in salience leads to a 1.9 increase in protest events. The effect becomes insignificant for salience values larger than 46, which corresponds to the top 17th percentile. The effect is robust to transformations of the dependent variable, as depicted in Figure 4.8 in Appendix 4.6.5. For these specifications, the effect is statistically significant at the 5% level for salience values between 20 and 60, which corresponds to 92% of all salience values.

By combining the estimated effects of Mechanisms 1 and 2, we derive an overall esti-

[17]The salience distribution is illustrated in Figure 4.9 in Appendix 4.6.5.

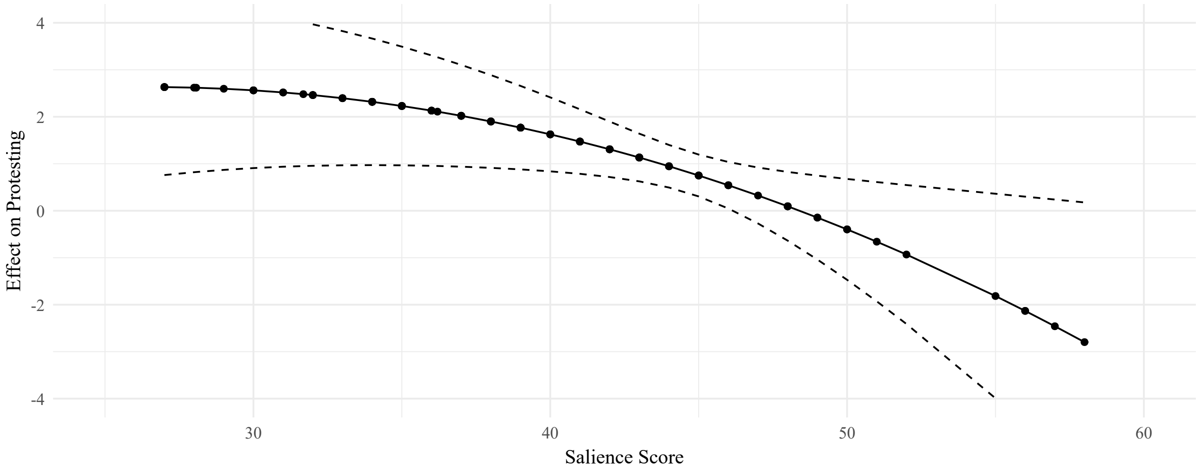


Figure 4.4: The graph illustrates the estimation results of the continuous salience treatment on protest events. The dashed lines indicate the 95% confidence interval. The salience score ranges from 20 to 60, which includes 92% of all salience values. Covariate Set 2 is applied.

mate for the entire channel, which we can compare to the results of the complete channel presented in Subsection 4.4.1. Specifically, taking the average coefficient of Mechanism 1 (3.02) and multiplying it by the density-weighted coefficient[18] of Mechanism 2 (2.21) yields a total effect of 6.69 (this number is calculated using the precise, unrounded values). This is about 72% of the total effect size of the average effect obtained in the previous subsection for the whole channel (9.3). Thus, we conclude that the salience mechanism is the dominant effect for short-run protest mobilization. In subsequent sections, we return to the question of what drives the remaining gap between the estimated indirect effect via salience and the total effect.

In conclusion, the results support Hypothesis 2 as they provide evidence for a salience channel from a context shock to protest outcomes. This effect holds for a variety of robustness tests. We continue this analysis in the following subsection, focusing on heterogeneity in the magnitude of context shocks and their implications for protest outcomes.

4.4.3 Differently Sized Context Shocks

To evaluate Hypothesis 3, which states that larger context shocks lead to stronger protest outcomes by increasing salience, we examine the estimation results presented in Table 4.4. We operationalize the magnitude of a context shock using the volume of media coverage associated with a specific fatal police encounter, measured by article counts. In this framework, greater media attention is expected to enhance salience, which, in turn,

[18]Only statistically significant values are considered, as we want to avoid extrapolating for high values of salience where almost no observed values are.

is predicted to lead to increased protest mobilization.

Distinguishing between context shocks, those that received significant media attention appear to have led to more protest events. This tendency is evident in the increasing coefficients for the regression outcomes in lines 1-3 in Table 4.4. For the regression adjustment specification, a small context shock increases protest by 9 units, a medium-sized shock by 14 units, and a large shock by 24.9 units on treated units. The size of the context shock, therefore, seems to substantially influence protest coordination. It should be noted that the estimates for these results are imprecisely estimated with relatively large standard errors. The salience shift appears to be highest for large context shocks, for which the salience level increases by 5.2-5.7 units. However, the coefficient sizes for small and medium context shocks are roughly the same, especially when accounting for the standard errors.

Table 4.4: Context Dependence: Context Shocks on Protest Outcomes

	RA	IPW	IPW RA	C. Match	P.S. Match
Dependent Variable: Protest Outcome					
Article C. ≥ 3	9.019*	6.358	6.280	7.083	3.883
	(4.858)	(5.139)	(5.082)	(5.125)	(5.913)
Observations	2905	2905	2905	2905	2905
Article C. ≥ 6	13.992*	10.201	9.936	12.414*	12.397*
	(7.171)	(7.531)	(7.518)	(7.194)	(6.974)
Observations	2875	2875	2875	2875	2875
Article C. ≥ 27	24.937*	21.277	21.317	21.214	18.250
	(14.063)	(14.198)	(14.420)	(14.372)	(14.789)
Observations	2845	2845	2845	2845	2845
Dependent Variable: Salience Level					
Article C. ≥ 3	3.443**	4.256***	4.547***	3.481**	3.345**
	(1.338)	(1.409)	(1.439)	(1.486)	(1.579)
Observations	2824	2824	2824	2824	2824
Article C. ≥ 6	2.998*	3.973**	4.184**	3.046*	2.793
	(1.639)	(1.679)	(1.682)	(1.769)	(2.165)
Observations	2794	2794	2794	2794	2794
Article C. ≥ 27	5.198**	5.659**	5.221**	5.155**	5.619***
	(2.279)	(2.422)	(2.564)	(2.567)	(1.555)
Observations	2764	2764	2764	2764	2764

Heteroskedasticity robust SE in parentheses; * p <0.1, ** p <0.05, *** p <0.01

Note: Estimation results of the Average Treatment Effect on the Treated (ATET) from the context shock on protest events and salience for the year 2022. The threshold values represent the top 75th, 50th, and 25th percentiles of the article count distribution. LASSO selection Covariate Set 1 is applied for the first 3 sets of estimators and Covariate Set 2 for the latter 3 sets.

The empirical findings are broadly consistent with the hypothesis, although they are not statistically significant. In both tables, we observe a positive association between the size of the context shock and the magnitude of its effect: larger context shocks tend to correspond to larger estimated coefficients for both protest activity and salience level. While overlapping confidence intervals of the point estimates prevent definitive conclusions about statistical differences between these coefficients, the overall pattern suggests a consistent relationship, particularly between high-salience events and mobilization.

Importantly, the article count is used to categorize context shocks based on the strength of the *external shock* on individuals' attention. If the analysis involved regressing article count on protest intensity or salience, the estimates would likely be biased because heightened protest activity or salience could drive increased media coverage. Instead, our approach involves excluding a subset of context shocks, specifically those that received little media attention, and examining whether the estimated treatment effects on protest incidence or salience differ from those of more widely covered shocks.

In sum, the results provide suggestive evidence that context shocks in the top quartile of media attention have the strongest effect on subsequent protest mobilization, with a substantial portion of this effect mediated through increased salience. The estimation of these effects is relatively imprecise, potentially due to the reduced number of context shocks at higher media-attention thresholds. The difference in response between small and medium-sized shocks remains uncertain.

4.4.4 Dynamic Effects

The final empirical estimation examines the dynamics of the salience mechanism. Following Hypothesis 4, a return of the context to its original value after a high-profile event should be accompanied by a decrease in protest mobilization. These dynamics are assessed by examining the publication gap between fatal police encounters and their media coverage, as well as a DLM that evaluates the intertemporal effects of context shocks on the salience level and protest outcomes.

The first piece of evidence is presented in Figure 4.5, which illustrates the timing of media coverage for fatal police encounters. Within the first 30 days, more than one-third of articles are published about these incidents. This indicates that the context, as defined in the model, is shocked most heavily in the immediate aftermath of a fatal police encounter. Over time, media interest in such incidents stabilizes at a low level until after four years, when more than 85% of the articles are published. The premise of Hypothesis 4, that the context normalizes over time, is therefore consistent with the data.

Assessing the timing of the salience mechanism, the presented evidence supports the

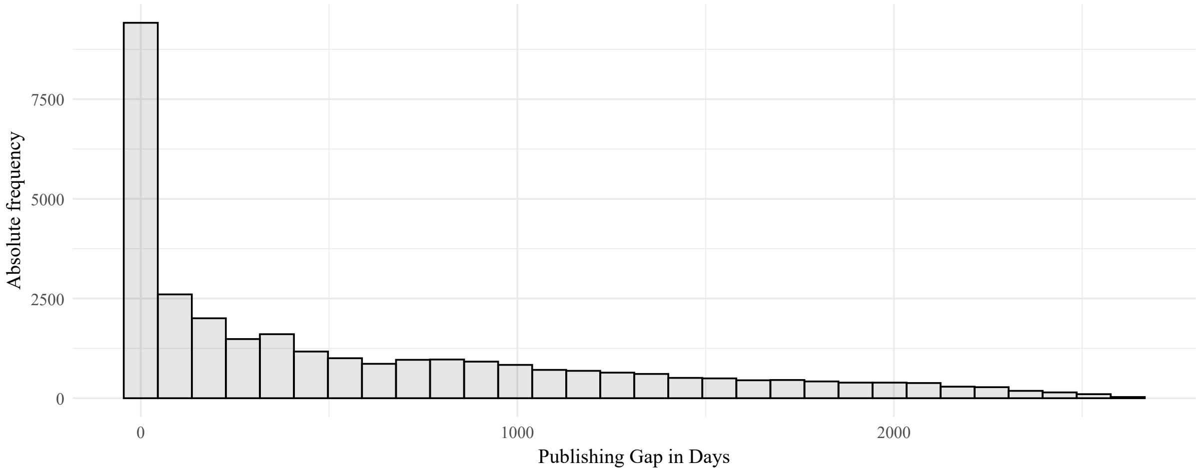


Figure 4.5: Publishing gap illustration as the difference in days between a fatal police encounter involving a Black individual and the publication of related articles. The analysis includes all incidents occurring between 2017 and 2022. Given this timeframe, the maximum observable publication lag for the most recent incidents is three years. 35.8% of articles are published within the first month after the fatal police encounter.

proposed mechanism's substantial explanatory power for short-term mobilization. To evaluate the persistence of salience-based protesting, we estimate a DLM, which is presented in Table 4.5. The contemporaneous effects identified in the previous regressions are replicated, as the context shock in 2022 significantly increases both the protest outcome and salience level. Regarding sustained mobilization, the context shocks from previous years continue to influence protest mobilization in the current period, but decrease in magnitude in the more precise second specification. The effect of context shocks on the salience level disappears after one period and even turns negative for one lag, speaking in favor of something like an "attention-hangover" after the treatment in 2020. We conclude that the salience mechanism plays a significant explanatory role in the short run, after which salience levels rapidly revert to their pre-treatment state. Consequently, the effect theorized in Hypothesis 4 materializes within the first year after a context shock, after which no statistically significant shift in salience is observed.

So, if the lagged effects of context shocks are not driven by the salience mechanism, what else drives these effects? A number of candidates are discussed in the literature. For one, networks may form during the early mobilization period. Thus, as soon as a context shock occurs, the salience mechanism mobilizes individuals to protests, where they form protest networks. Such networks may be capable of sustaining protest mobilization within a smaller circle of individuals without relying on increased attention in the larger regional environment. This mechanism aligns with the findings of Bursztyn *et al.* (2021), who show

Table 4.5: Distributed Lag Model: Context Shocks on Protest Outcomes.

Dependent Variable	*Protest Events*		*Salience Level*	
	Level	Inverse hyp.	County Level	Regional Level
Context Shock 2018	5.421	0.202**	1.155	1.715
	(3.297)	(0.094)	(0.795)	(1.757)
Context Shock 2019	3.559	0.267***	-0.913	-1.988
	(2.571)	(0.090)	(0.815)	(1.807)
Context Shock 2020	4.603	0.324***	-1.829**	-5.479**
	(2.809)	(0.096)	(0.725)	(2.384)
Context Shock 2021	5.772*	0.344***	0.091	1.514
	(3.266)	(0.096)	(0.769)	(1.695)
Context Shock 2022	3.473*	0.399***	2.751***	6.631***
	(2.075)	(0.091)	(0.865)	(1.705)
LRP	22.828**	1.536***	1.255	2.393
	(8.977)	(0.159)	(1.274)	(2.084)
Observations	3007	3007	2924	201

Heteroskedasticity robust SE in parentheses; * p <0.1, ** p <0.05, *** p <0.01

Note: Distributed Lag Model for the effects of past and current context shocks on protest outcomes and salience level in 2022. LRP denotes the long-run elasticity. Covariate Set 1 is applied for protest events and Covariate Set 2 for salience as dependent variable.

that initial protest participation, when reinforced by social network effects within close peer groups, can generate sustained mobilization over time. A complementary mechanism may be that organizational capacities were built during the early protest stage, facilitating later mobilization, as networks could.

In summary, we presented a variety of empirical evidence supporting the existence of a salience mechanism. This mechanism has immediate effects and socio-economic significance. However, regarding sustained mobilization in the aftermath of high-profile events, the explanatory power of the proposed mechanism is low, indicating that other approaches are needed for explanations. Given the current dynamics in attention markets, such as social media, this finding should not be too surprising. The only thing faster than the emergence of hype is its inevitable end, which also seems to be true for attention-motivated protesting.

4.4.5 Robustness

The regression results in the previous subsection draw a clear picture of the proposed salience mechanism from a context shock to increased protesting via salience. In this sensitivity analysis, we present supplementary empirical evidence to demonstrate the robustness of the salience mechanism. To this end, we estimate robustness checks for the period between 2018 and 2022.

To demonstrate the robustness of the salience mechanism across different years, we assess the binary treatment model and estimate it for the years 2018, 2020, and 2022. Generally, the effect of context shocks on protest outcomes varies substantially across these years, as the estimates in Table 4.11 in Appendix 4.6.5 illustrate. A context shock is found to increase protest events in treated counties by 18.1 to 25 units in 2020. This effect is smaller for 2022; here, a context shock increases protest events by 8.5-10.5 units. The smallest effects are identified for 2018, where the context shock led to an increase of 0.5-0.6 units only. The results, therefore, indicate a much greater sensitivity of protest outcomes to context shocks since 2020, which likely is the result of an interaction effect between the context shock and the murder of George Floyd in the same year.

Besides the results presented in Table 4.11 in Appendix 4.6.5, there are further robustness checks presented in Appendix 4.6.5 and 4.6.5. In Table 4.12 in Appendix 4.6.5, the baseline regressions for 2019 and 2021 generally show mixed results, with some specifications in line with the previous findings, while others lack statistical identification. For the years 2018 and 2020, we run the complete set of robustness checks (Table 4.12 in Appendix 4.6.5). The results generally demonstrate the robustness of the previous findings. For 2018, the specification where we trimmed extreme propensity scores loses significance, potentially due to smaller statistical power as a consequence of fewer observations and a generally small point estimate that we find prior to 2020. For 2020, the extensive margin is not statistically identified, potentially due to the widespread mobilization of protests during this year. As many counties showed protests anyway during this year, it is likely to become infeasible to identify the extensive margin for this year. However, the eight alternative robustness checks for both years are in line with the previous results and thereby demonstrate substantial robustness of the empirical identification. The effect of salience on protest outcomes is estimated using the same continuous treatment model as in Subsection 4.4.2 and presented in Figure 4.10 in Appendix 4.6.5. The effect size of salience on protests is generally much higher in 2020 compared to 2022 and decreases in size for 2018, where the effect is barely statistically significant.

4.5 Conclusion

In this paper, we develop an experience-based salience model that incorporates the attentional context in protest decisions. In this framework, high-profile events such as fatal police encounters with Black individuals act as context shocks by altering the perception of protests and drawing attention to the benefits of protesting, thereby increasing their salience. This shifted salience can change protest decisions and increase participation, even when the objective benefits or costs remain unchanged. We assess the model's predictions using a set of treatment evaluation methods in the context of BLM and anti-police violence protests in the US. Overall, we find that fatal police encounters significantly affect protest outcomes, with a large portion of this effect mediated through salience. We distinguish between the sizes of the fatal police encounters by categorizing them according to their article count and evaluating their intertemporal dynamics using a DLM. The results indicate that larger context shocks exert a stronger influence on protest mobilization than smaller ones. Moreover, attention-driven protest activity tends to emerge rapidly following such shocks but dissipates within approximately one year.

The paper contributes to a rich literature on protest models. The attentional channel has the potential to be incorporated into many of the previous protest models. For instance, incorporating critical events and their attentional consequences into a threshold model or a model in which protests are strategic complements could help to explain fast-paced protest mobilization. The salience mechanism should be understood as complementary to previous protest models, as it explicitly incorporates subconscious perceptual aspects of protest mobilization.

One promising path for future research is to more thoroughly assess the dynamic aspects of attention-based protesting. The results of this study indicate that attention-motivated protesting is a short-term phenomenon that cannot fully explain sustained protesting. Other factors, such as the formation of networks, organizational capacities, or information cascades, are therefore more likely to shed light on intertemporal protest dynamics. A broader perspective that combines explanatory elements from multiple theories may offer a more accurate and precise account of intertemporal protest dynamics.

4.6 Appendix

4.6.1 Mathematical Appendix

Proof of Lemma 1: For discrete-time $s \in \{0, ..., t\}$ the reference benefit at time t is given by

$$\bar{b}_t = \sum_{b_s \in M_t} b_s \cdot w(e_s, e_t) = \sum_{s=0}^{t} b_s \cdot w(e_s, e_t).$$

Then it holds $\bar{b}_t > (<)\ \bar{b}_{t-1}$ if and only if

$$\begin{aligned}
& \frac{\sum_{s=0}^{t} S(e_s, e_t) \cdot b_s}{\sum_{s=0}^{t} S(e_s, e_t)} - \bar{b}_{t-1} && \gtrless 0 \\
\Leftrightarrow\quad & \frac{\sum_{s=0}^{t} S(e_s, e_t) \cdot b_s - \bar{b}_{t-1} \cdot \sum_{s=0}^{t} S(e_s, e_t)}{\sum_{s=0}^{t} S(e_s, e_t)} && \gtrless 0 \\
\Leftrightarrow\quad & \frac{\sum_{s=0}^{t} S(e_s, e_t) \cdot [b_s - \bar{b}_{t-1}]}{\sum_{s=0}^{t} S(e_s, e_t)} && \gtrless 0 \\
\Leftrightarrow\quad & \sum_{s=0}^{t} S(e_s, e_t) \cdot [b_s - \bar{b}_{t-1}] && \gtrless 0
\end{aligned}$$

As $f(k_s) = b_s$ is a continuous and injective function, a global context shock k_t leads to the perceived benefit b_t reaching a new extreme level, i.e., $b_t > max_{s<t} b_s$ or $b_t < min_{s<t} b_s$. Thus, if k_t induces a new maximum (minimum) level it holds

$$S(e_i, e_t) > S(e_j, e_t) \Leftrightarrow b_i > (<)\ b_j, \quad \forall i, j \in \{0, ..., t\}$$

since $S(e_s, e_t)$ is strictly decreasing in $|k_t - k_s|$. Then, by Chebyshev's sum inequality, it holds

$$\begin{aligned}
& t \cdot \sum_{s=0}^{t} S(e_s, e_t) \cdot [b_s - \bar{b}_{t-1}] \gtrless \left(\sum_{s=0}^{t} S(e_s, e_t) \right) \cdot \left(\sum_{s=0}^{t} b_s - \bar{b}_{t-1} \right) \\
\Leftrightarrow\quad & \frac{\sum_{s=0}^{t} S(e_s, e_t) \cdot [b_s - \bar{b}_{t-1}]}{\sum_{s=0}^{t} S(e_s, e_t)} \gtrless \frac{1}{t} \cdot \sum_{s=0}^{t} b_s - \bar{b}_{t-1} \\
\Leftrightarrow\quad & \frac{\sum_{s=0}^{t} S(e_s, e_t) \cdot b_s}{\sum_{s=0}^{t} S(e_s, e_t)} - \bar{b}_{t-1} \cdot \frac{\sum_{s=0}^{t} S(e_s, e_t)}{\sum_{s=0}^{t} S(e_s, e_t)} \gtrless \frac{1}{t} \cdot \sum_{s=0}^{t} b_s - \frac{1}{t} \cdot \sum_{s=0}^{t} \bar{b}_{t-1} \\
\Leftrightarrow\quad & \frac{\sum_{s=0}^{t} S(e_s, e_t) \cdot b_s}{\sum_{s=0}^{t} S(e_s, e_t)} \gtrless \frac{1}{t} \cdot \sum_{s=0}^{t} b_s
\end{aligned}$$

such that for $\bar{b}_{t-1} \leq (\geq)\ \frac{1}{t} \cdot \sum_{s=0}^{t} b_s$, it always holds $\bar{b}_t > (<)\ \bar{b}_{t-1}$ and Condition (i) is shown.

In general, Inequality 4.10 is always fulfilled if

$$\alpha_i \geq \alpha_j \Leftrightarrow b_i \geq (\leq)\ b_j, \quad \forall i,j \in \{0,...,t-1\}$$

with $\alpha_s = \frac{S(e_s,e_t)}{S(e_s,e_{t-1})}$ since then, by the weighted Chebyshev's sum inequality, it holds

$$\begin{aligned}
& \frac{\sum_{s=0}^{t-1} \alpha_s \cdot S(e_s,e_{t-1}) \cdot [b_s - \bar{b}_{t-1}]}{\sum_{s=0}^{t-1} S(e_s,e_{t-1})} \underset{\leq}{\geq} \left(\frac{\sum_{s=0}^{t-1} \alpha_s \cdot S(e_s,e_{t-1})}{\sum_{s=0}^{t-1} S(e_s,e_{t-1})} \right) \cdot \left(\frac{\sum_{s=0}^{t-1} S(e_s,e_{t-1}) \cdot [b_s - \bar{b}_{t-1}]}{\sum_{s=0}^{t-1} S(e_s,e_{t-1})} \right) \\
\Leftrightarrow \quad & \sum_{s=0}^{t-1} \alpha_s \cdot S(e_s,e_{t-1}) \cdot [b_s - \bar{b}_{t-1}] \underset{\leq}{\geq} \left(\frac{\sum_{s=0}^{t-1} \alpha_s \cdot S(e_s,e_{t-1})}{\sum_{s=0}^{t-1} S(e_s,e_{t-1})} \right) \cdot \left(\sum_{s=0}^{t-1} S(e_s,e_{t-1}) \cdot [b_s - \bar{b}_{t-1}] \right) \\
\Leftrightarrow \quad & \sum_{s=0}^{t-1} S(e_s,e_t) \cdot [b_s - \bar{b}_{t-1}] \underset{\leq}{\geq} 0
\end{aligned}$$

The last equivalence transformation holds since we have

$$\begin{aligned}
& \bar{b}_{t-1} = \frac{\sum_{s=0}^{t-1} S(e_s,e_{t-1}) \cdot b_s}{\sum_{s=0}^{t-1} S(e_s,e_{t-1})} \\
\Leftrightarrow \quad & \frac{\sum_{s=0}^{t-1} S(e_s,e_{t-1}) \cdot b_s - \bar{b}_{t-1} \cdot \sum_{s=0}^{t-1} S(e_s,e_{t-1})}{\sum_{s=0}^{t-1} S(e_s,e_{t-1})} = 0 \\
\Leftrightarrow \quad & \sum_{s=0}^{t-1} S(e_s,e_{t-1}) \cdot [b_s - \bar{b}_{t-1}] = 0
\end{aligned}$$

As $b_t > (<)\ \bar{b}_{t-1}$, Inequality 4.10 is fulfilled and Condition (ii) is shown.

q.e.d

4.6.2 Beta Densities

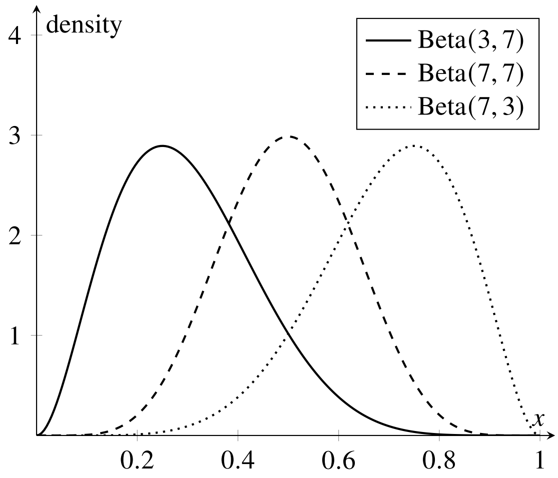


Figure 4.6: Densities for different $Beta(\alpha, \beta)$ distributions.

4.6.3 LASSO Double Selection

LASSO Double Selection

In order to obtain causal estimates, the estimation strategy relies on the selection-on-observables assumption (or conditional independence assumption). This assumption holds that after controlling for a set of covariates, the selection into treatment is random and, thereby, not endogenous to the outcome. One approach for justifying this assumption is to include a large set of covariates that are frequently used in the literature. However, this procedure has its downsides as it is subject to researchers' prior conceptions of fitting control variables and may lead to inflated standard errors due to multicollinearity.

LASSO (least absolute shrinkage and selection operator) double selection is a suitable method for establishing the selection-on-observables assumption in a way that avoids subjectivity and provides valid test statistics for causal inference (Belloni *et al.*, 2014; Huber, 2023). Subjective perceptions about suitable control variables have a smaller influence as the researcher only selects a large set of potential covariates. The necessary control variables for establishing the selection-on-observables assumption in the causal analysis are selected from this set in a data-driven way, reducing the researcher's involvement. Additionally, LASSO double selection increases the confidence in a model's test statistics. Treatment effect estimators (and ordinary regression methods) are generally not robust to human-made rules for variable selection, potentially leading to invalid test statistics (Huber, 2023). Procedures, such as LASSO double selection, alleviate such concerns, as the inclusion of covariates is determined in a data-driven way. These advantages, of course, only materialize if the set of potential control variables is sufficiently large. To ensure this, 76 potential covariates are selected in this study as presented in Table 4.6 and 4.7 in Appendix 4.6.3.

Following Belloni *et al.* (2014), the selection procedure and the post-LASSO estimations rely on two assumptions: Neyman-orthogonality and approximate sparsity. Neyman-orthogonality (Neyman, 1959) ensures that the estimation of treatment effects is robust to approximation errors in the nuisance parameters (Huber, 2023). That is, the target parameter is insensitive to perturbations in the LASSO estimation due to regularization. Furthermore, approximate sparsity ensures that the number of important covariates is small relative to the sample size n. Given these assumptions, a large set of variables can be reduced using LASSO double selection, and causal analysis may be conducted via post-LASSO OLS estimation. This procedure, thanks to the given assumptions, is $\sqrt{n}$-consistent.

LASSO double selection relies on the LASSO regularization technique stemming from the predictive machine learning literature (James *et al.*, 2013). The model is presented in Equation 4.20. The left-hand side of the curly bracket is the OLS minimization problem for which the squared residuals are minimized. The right-hand side introduces a penalty term for absolute coefficient size. The penalty term introduces (regularization-) bias and reduces variance. The tuning parameter λ is chosen to balance bias and variance so as to maximize predictive performance in the validation set. All coefficients that do not contribute to the predictive performance of the model (reducing the OLS term) are shrunken to 0 (to reduce the penalty term).

$$\text{minimize}_{\boldsymbol{\beta}} \left\{ \sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{j=1}^{p} \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^{p} |\beta_j| \right\} \tag{4.20}$$

Causal inference with LASSO double selection is conducted in three steps (Belloni *et al.*, 2014). The first is a LASSO regression of the outcome variable on a constant and all potential covariates. The second is a LASSO regression of the treatment variable on a constant and all potential covariates. Each variable with a positive absolute coefficient in one of these LASSO regressions is included in a list of selected control variables. Finally, a causal OLS regression of the outcome variable on a constant, the treatment variable, and all selected variables is run. The treatment variable's coefficient is causally interpretable as all potentially confounding variance is controlled through the selected control variables.

LASSO double selection is applied to the data at hand. For the first empirical specification, two LASSO regressions are run: (1) protest events on a constant and all potential control variables, and (2) focal points on a constant and all potential control variables. The potential control variables are summarized in Table 4.6 and 4.7. All variables with a positive absolute coefficient in at least one of these regressions are included as control variables for the different estimation techniques in the following. The LASSO estimation is conducted in R using the *High-Dimensional Metrics* package (Chernozhukov *et al.*, 2016).

Included Variables for LASSO Selection Procedure

Table 4.6: Included Variables for LASSO Selection Procedure (1/2)

Variable	Description
population	Population number
log population	Logarithm population number
ages 0 19	Population share 0 to 19 year olds
ages 20 29	Population share 20 to 29 year olds
ages 30 39	Population share 30 to 39 year olds
ages 40 49	Population share 40 to 49 year olds
ages 50 59	Population share 50 to 59 year olds
ages 60 69	Population share 60 to 69 year olds
ages 70 79	Population share 70 to 79 year olds
ages 80 older	Population share 80 years and older
white	Polulation share white individuals
black	Population share black individuals
native	Population share native individuals
asian	Population share asian individuals
native pacific	Population share native pacific individuals
other	Population share other race individuals
two or more	Population share two or more race individuals
average household size	Average household size
married	Population share of married individuals
not married	Population share of not married individuals
highschool degree	Population share highschool degree
no degree	Population share no educational degree
university degree	Population share university degree
employed	Employment rate
unemployed	Unemployment rate
employed youth	Employment rate for 16 to 24 year olds
unemployed youth	Unemployment rate for 16 to 24 year olds
k10	Population share with income of 0 to 10,000 USD.
k30	Population share with income of 20,000 to 30,000 USD.
k40	Population share with income of 30,000 to 40,000 USD.
k50	Population share with income of 40,000 to 50,000 USD.
k60	Population share with income of 50,000 to 60,000 USD.
k75	Population share with income of 60,000 to 75,000 USD.
k100	Population share with income of 75,000 to 100,000 USD.
k125	Population share with income of 100,000 to 125,000 USD.
k150	Population share with income of 125,000 to 150,000 USD.
k20	Population share with income of 10,000 to 20,000 USD.
k200	Population share with income of 150,000 to 200,000 USD.
kmore200	Population share with income of more than 200,000 USD.

Note: Set of potential control variables, the final set of control variables is determined by LASSO double selection. The variables are retrieved from the US Census if not indicated otherwise.

Table 4.7: Included Variables for LASSO Selection Procedure (2/2)

Variable	Description
median household income	Median household income
log median household income	Logarithm median household income
median household income white	Median household income white individuals
log median HI white	Logarithm median household income white individuals
median household income black	Median household income black individuals
log median HI black	Logarithm median household income black individuals
median household income race gap	Median household income race gap
no other income	Population share with labor income only
other income	Population share with non-wage income
per capita income	Income per capita
log per capita income	Logarithm of income per capita
gini	Gini coefficient
rent percent householde income	Rent as percentage of household income
median gross rent	Median gross rent
log median gross rent	Logarithm median gross rent
im 1990 1999	Share of population of immigrants that came to the US beween 1990 and 1999
im 2000 2009	Share of population of immigrants that came to the US beween 2000 and 2009
im 2010 later	Share of population of immigrants that came to the US beween 2010 and later
im before 1990	Share of population of immigrants that came to the US before 1990
health insurance	Share of population with health insurance
no health insurance	Sahre of population without health insurance
election clos	Votes 2nd Cand. / Votes Winner (MIT Election Data and Science Lab, 2018)
pol fractionalization ELF	Political fractionalization ELF score of voter groups (MIT Election Data and Science Lab, 2018)
election clos poppoll	Votes 2nd cand. / votes winner (MIT Election Data and Science Lab, 2018)
election clos states poppoll	Votes 2nd cand. / votes winner (states) (MIT Election Data and Science Lab, 2018)
election clos counties poppoll	Votes 2nd cand. / votes winner (counties) (MIT Election Data and Science Lab, 2018)
online network	Number of Facebook connections/ possible number of Facebook connections (logarithmized, 2021 only) Bailey *et al.* (2018)
family unity	Weighted measure of births to unwed women, children living in single-parent households, and percentage of women aged 35-44 who are married (Congress, 2018)
community health	Index containing share of (1) adults who volunteered, (2) discussed community affairs, or/and (3) worked for improving the community (Congress, 2018)
institutional health	Adherence to state institution measure (Congress, 2018)
social capital	Indicator composed of different dimensions of social capital like trust, networks, or shared values (Congress, 2018)
collective efficacy	Violent crime rate (Congress, 2018)
charitable organizations	Number of charitable organizations per capita (Lecy, 2023)
area	Area of land
density	Population number divided by area size (own calculation)
center 1	Dummy variable that equals 1 if density $>$ 40,000 (own calculation)
center 2	Dummy variable that equals 1 if density $>$ 20,000 (own calculation)

Note: Set of potential control variables, the final set of control variables is determined by LASSO double selection. The variables are retrieved from the US Census if not indicated otherwise.

LASSO Selected Covariate Sets

The following are covariate sets that are obtained by applying the LASSO double selection procedure. The notes below all regression tables indicate which of the covariate sets are used for the specific estimations.

Covariate Set 1: log population, the share of age group 30-39, the share of white individuals, the share of individuals with university degree, the Gini coefficient, the share of immigrants (from after 2010), binary variable whether democrats have majority support in the county, and a measure of collective efficacy.

Covariate Set 2: log population, the share of age group 30-39, the share of age group 50-59, the share of white individuals, share with a university degree, the Gini coefficient, the share of immigrants (from after 2010), the share of immigrants (from after 1990), share of individuals without health insurance, binary variable whether democrats have majority support in the county, measure of political fractionalization, measure of institutional health, and a measure of collective efficacy.

Covariate Set 3: log population, the share of age group 30-39, the share of white individuals, share of married individuals, share with a high school degree, share with a university degree, the share of individuals with a yearly income over 200,000$, the Gini coefficient, the share of immigrants (immigrated after 2010), binary variable whether democrats have majority support in the county, a measure of collective efficacy, and the number of charitable organizations per capita.

Covariate Set 4: log population, the share of age group 30-39, the share of white individuals, share of married individuals, share with a high school degree, share with a university degree, the share of individuals with a yearly income over 200,000$, the Gini coefficient, the share of immigrants (immigrated after 2010), binary variable whether democrats have majority support in the county, a measure of collective efficacy, and the number of charitable organizations per capita, share of married individuals, share with a high school degree, the share of individuals with a yearly income over 200,000$, and the number of charitable organizations per capita.

Covariate Set 5: log population, the share of age group 30-39, the share of white individuals, share with a university degree, the Gini coefficient, the share of immigrants (immigrated after 2010), the share of immigrants (immigrated after 1990), binary variable whether democrats have majority support in the county, and a measure of collective efficacy.

4.6.4 Binary Treatment

Overlap Assumption

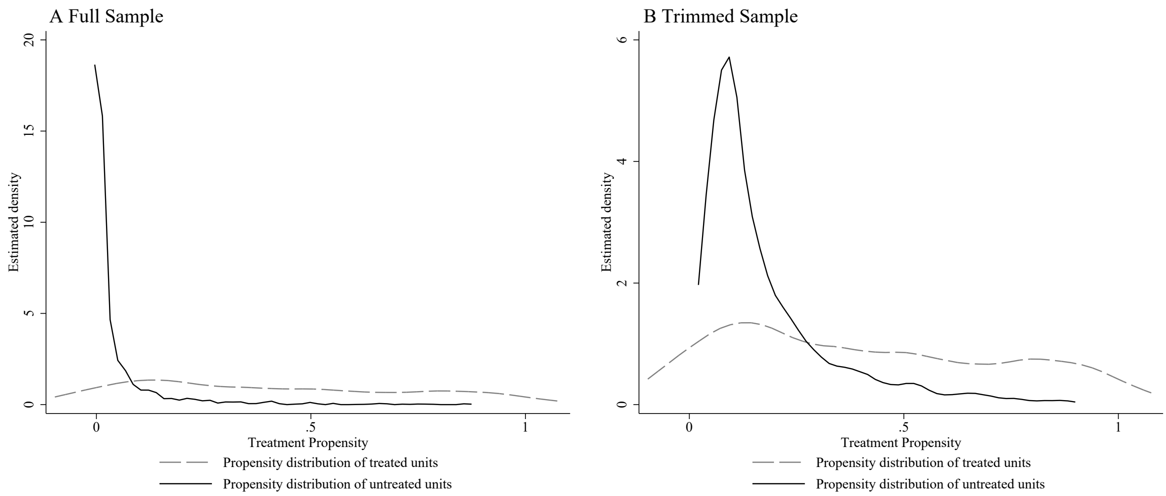


Figure 4.7: Smoothed density of the propensity score for the treatment and control group. B presents a trimmed sample in which all propensity scores below 0.05 were dropped for the control group to make the overlap of the remaining units visible. The smoothing is conducted relying on the Epanechnikov kernel.

Estimator Description

(1) A regression adjustment (RA) estimator is the first assessment of the target parameter from Equation 4.17 (Wooldridge, 2010). The estimator is depicted in Equation 4.21 where w_i is a dummy equal to 1 for all treated units and $\hat{m}$ is a linear model predicting the unobserved outcome. The estimator computes the average difference between the treated outcome and the hypothetical untreated outcome of all treated units in the sample. The hypothetical outcome is a predicted value from the linear model.

$$\hat{\tau}_{\mathrm{RA}} = \left(\sum_{i=1}^{N} W_i\right)^{-1} \left\{\sum_{i=1}^{N} W_i \left[\hat{\boldsymbol{m}}_1\left(\mathbf{X}_i\right) - \hat{\boldsymbol{m}}_0\left(\mathbf{X}_i\right)\right]\right\} \tag{4.21}$$

(2) The second static estimator is an inverse probability weighting (IPW) estimator that weights observations differently depending on how likely it is that the observation had the opposite treatment status. The estimator is presented in Equation 4.22 (Wooldridge, 2010). Observations closer to the counterfactual group are weighted more strongly in the estimation. In the first stage, the propensity score function $P(X)$ is obtained using a logit model. In the second stage, the observations are weighted by the inverse of the propensity score. Given the obtained inverse propensity scores, the ATET is estimated according to Equation 4.22.

$$\widehat{\tau}_{\mathrm{IPW}} = N^{-1} \sum_{i=1}^{N} \frac{\left[W_i - P(X_i)\right] Y_i}{\frac{N_1}{N}\left[1 - P(X_i)\right]} \tag{4.22}$$

(3) Following Wooldridge (2007), inverse probability-weighting can be enhanced by using a regression adjustment technique, making the estimator doubly-robust. In the procedure, the inverse probability weights are computed using a probit model as for IPW estimation. Given the inverse-probability weights, a weighted regression is fit for each treatment status as in the RA estimation. This regression is used to obtain the hypothetical outcomes, whose average differences are the ATE, and the average differences among the treated are the ATET of interest. The inverse probability-weighting regression adjustment (IPWRA) estimator thus combines the estimators from Equation 4.21 and 4.22.

(4) A fourth estimation strategy is propensity score matching (PSM) depicted in Equation 4.23 and 4.24. In the first step, a propensity score for the likelihood of treatment is calculated using a probit model and the selected covariates, resulting in $P(X)$. Conditioning on this probability is equivalent to controlling for the covariates themselves (Rosenbaum and Rubin, 1983). Instead of controlling for the propensity score in an OLS regression, a matching estimator is applied. This matching estimator matches the three nearest untreated observations in terms of the difference in the propensity score to a treated observation and compares the differences in the outcome variable.

$$\widehat{\tau}_{\text{PSM}} = \frac{1}{n_1} \sum_{i=1}^{n} W_i \left(Y_i - Y_{j(i)} \right) \tag{4.23}$$

where

$$Y_{j(i)} = \sum_{j:D_j=0} I \left(|P(X_j) - P(X_i)| = \min_{l:D_l=0} |P(X_l) - P(X_i)| \right) Y_j \tag{4.24}$$

(5) A final estimation strategy is direct covariate matching (COM). Instead of using the propensity score, all covariates are used to find the three nearest untreated neighbors of each treated observation. Importantly, all covariates are standardized. The nearest units are determined by the Mahalanobis distance between observations – this distance measure accounts for all differences between the standardized selected covariates while controlling for the correlation among them (Huber, 2023). The ATET estimator is formally depicted in Equation 4.25 and 4.26 where $\|X_l - X_i\|$ denotes the Mahalanobis distance between observation l and i.

$$\widehat{\tau}_{\text{COM}} = \frac{1}{n_1} \sum_{i=1}^{n} W_i \left(Y_i - Y_{j(i)} \right) \tag{4.25}$$

where

$$Y_{j(i)} = \sum_{j:D_j=0} I \left(\|X_j - X_i\| = \min_{l:D_l=0} \|X_l - X_i\| \right) Y_j \tag{4.26}$$

Covariate Balance and Placebo Tests

Table 4.8: Covariate Balance and Placebo Tests

SMD	RA	IPW	IPW RA	C. Match	P.S. Match
Log Population		0.119	0.033	0.033	-0.045
Age Group 30-39		0.129	0.054	0.054	-0.055
Share White Ind.		-0.072	-0.026	-0.026	-0.047
Share Uni. Degree		0.057	0.001	0.001	-0.059
Gini Coefficient		0.166	0.138	0.138	0.184
Share Immig. (a. 2010)		0.012	-0.039	-0.039	-0.119
Democrat Majority		0.097	0.047	0.047	0.036
Collective Efficacy		-0.209	-0.201	-0.201	-0.225
Overidentification		0.8011	0.8943		
Placebo Treatment	14.123***	10.223*	9.280*	12.518**	9.867*
(2021)	(5.125)	(5.356)	(5.416)	(5.180)	(5.137)
Placebo Treatment	21.433***	10.978	7.843	17.186***	6.893
(2020)	(5.402)	(6.974)	(7.802)	(5.377)	(7.015)
Placebo Treatment	0.654**	0.317	0.185	0.563*	0.102
(2019)	(0.325)	(0.356)	(0.374)	(0.327)	(0.327)
Observations	3007	3007	3007	3007	3007

Heteroskedasticity robust SE in parentheses; * p <0.1, ** p <0.05, *** p <0.01

Note: Covariate balance and placebo tests. The first rows present the standardized mean differences (SMD) of the control vs. treatment group after weighting and matching. Covariate balance is also tested using the overidentification test proposed by Imai and Ratkovic (2014) for weighting estimators. For the placebo tests, the protest outcome variable of the baseline specification is replaced with values for 2021, 2020, and 2019.

4.6.5 Sensitivity Analysis

Robustness Static Model

Table 4.9: Robustness Binary Treatment Evaluation: Protest Outcome

	RA	IPW	IPW RA	C. Match	P.S. Match
Inverse hyper. Y	0.734***	0.469***	0.401***	0.648***	0.244
	(0.110)	(0.133)	(0.141)	(0.126)	(0.155)
Observations	2968	2968	2968	2968	2968
Binary Y	0.242***	0.129***	0.105**	0.193***	0.058
	(0.034)	(0.041)	(0.043)	(0.046)	(0.036)
Observations	3007	3007	3007	3007	3007
Trimmed Sample	11.847***	10.905**	10.647**	12.249***	10.232**
(75 mi spillovers)	(4.302)	(4.394)	(4.437)	(4.280)	(4.411)
Observations	1345	1345	1345	1345	1345
Trimmed Sample	11.649***	11.177**	10.946**	12.586***	10.981**
(100 mi spillovers)	(4.338)	(4.425)	(4.471)	(4.286)	(4.346)
Observations	966	966	966	966	966
Trimmed Sample	5.036	5.985	5.990	8.163**	6.060
(overlap)	(4.137)	(3.750)	(3.730)	(3.602)	(3.882)
Observations	714	714	714	714	714

Heteroskedasticity robust SE in parentheses; * p <0.1, ** p <0.05, *** p <0.01

Note: Robustness estimation results of the Average Treatment Effect on the Treated (ATET) from the context shock on protest events for the year 2022. For the inverse hyperbolic Y Covariate Set 4 and the binary Y Covariate Set 3 are applied. All other estimations are based on Covariate Set 1.

Table 4.10: Robustness Binary Treatment Evaluation: Salience Level

	RA	IPW	IPW RA	C. Match	P.S. Match
Inverse hyper. Y	0.043**	0.072***	0.073***	0.066***	0.086***
	(0.017)	(0.019)	(0.020)	(0.020)	(0.027)
Observations	2946	2946	2946	2946	2946
Trimmed Sample	3.819***	6.653***	7.144***	5.263***	8.938***
(75 mi spillovers)	(1.390)	(1.488)	(1.571)	(1.263)	(1.693)
Observations	1232	1232	1232	1232	1232
Trimmed Sample	3.632**	6.368***	6.574***	5.395***	8.094***
(100 mi spillovers)	(1.648)	(2.001)	(2.186)	(1.317)	(1.729)
Observations	859	859	859	859	859
Trimmed Sample	4.489***	4.032***	4.072***	2.888***	3.659***
(overlap)	(1.001)	(0.967)	(0.985)	(0.891)	(1.154)
Observations	683	683	683	683	683
Higher Level	5.363***	6.857***	6.835***	4.346***	5.886***
	(1.551)	(1.870)	(1.873)	(1.414)	(1.087)
Observations	201	201	201	201	201

Heteroskedasticity robust SE in parentheses; * p <0.1, ** p <0.05, *** p <0.01

Note: Robustness estimation results of the Average Treatment Effect on the Treated (ATET) from the context shock on salience for the year 2022. For the inverse hyperbolic Y Covariate Set 5 is applied. All other estimations are based on Covariate Set 2.

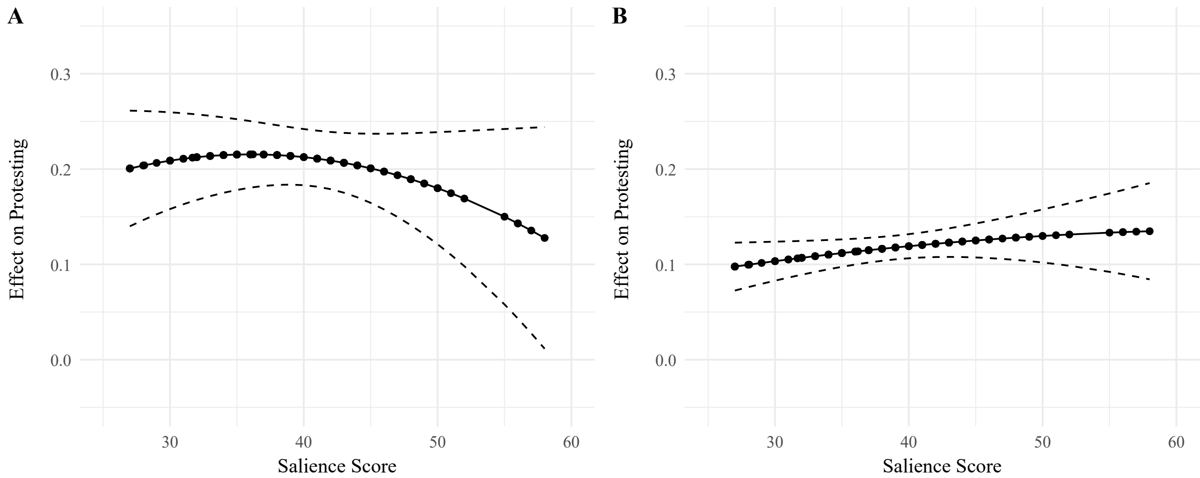


Figure 4.8: Robustness: Results of continuous treatment of salience on protest outcome. The dashed lines indicate the 95% confidence interval. Graph A depicts the fuzzy treatment effect relying on the inverse hyperbolic sine transformed protest variable as the outcome, while Graph B illustrates the fuzzy effect for a binary protest variable as the outcome. LASSO selection Covariate Set 2 is applied.

Distribution Salience

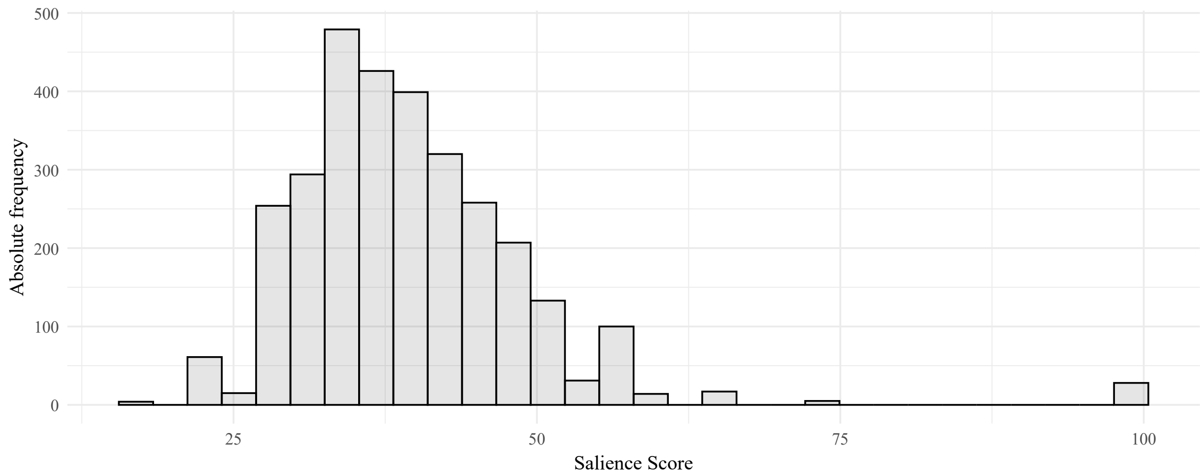


Figure 4.9: Distribution salience scores in 2022 based on Google Trends data.

Robustness Year Variation

Table 4.11: Context Dependence: Context Shocks on Protest Outcomes

	RA	IPW	IPW RA	C. Match	P.S. Match
Dependent Variable: Protest Outcome					
Context Shock 2018	0.620*	0.557*	0.530	0.526	0.496
	(0.327)	(0.331)	(0.334)	(0.328)	(0.337)
Observations	3014	3014	3014	3014	3014
Context Shock 2020	25.019***	20.213***	18.358***	21.920***	18.146***
	(6.075)	(6.269)	(6.305)	(5.883)	(7.028)
Observations	3013	3013	3013	3013	3013
Context Shock 2022	10.463**	8.768*	8.511*	9.044**	9.728**
	(4.364)	(4.474)	(4.456)	(4.525)	(4.292)
Observations	3007	3007	3007	3007	3007
Dependent Variable: Salience Level					
Context Shock 2018	0.130*	0.225***	0.227***	0.246***	0.276***
	(0.076)	(0.079)	(0.083)	(0.076)	(0.086)
Observations	2926	2926	2926	2926	2926
Context Shock 2020	0.100	0.148**	0.148*	0.114	0.139
	(0.069)	(0.072)	(0.077)	(0.070)	(0.091)
Observations	2934	2934	2934	2934	2934
Context Shock 2022	0.157***	0.172***	0.170**	0.176***	0.154***
	(0.059)	(0.064)	(0.066)	(0.058)	(0.050)
Observations	2907	2907	2907	2907	2907

Heteroskedasticity robust SE in parentheses; * p <0.1, ** p <0.05, *** p <0.01

Note: Estimation results of the Average Treatment Effect on the Treated (ATET) from the context shock on protest events and salience for the year 2022. LASSO selection Covariate Set 1 is applied for the first 3 sets of estimators and Covariate Set 2 for the latter 3 sets.

Table 4.12: Robustness Year Variation

	RA	IPW	IPW RA	C. Match	P.S. Match
Shooting 2019	0.792**	0.509	0.357	0.760**	0.439
	(0.367)	(0.396)	(0.422)	(0.360)	(0.384)
Observations	3014	3014	3014	3014	3014
Shooting 2021	13.096**	3.361	-0.827	12.694**	0.525
	(5.474)	(9.093)	(11.118)	(5.089)	(8.786)
Observations	3013	3013	3013	3013	3013
Inverse hyper. Y 2018	0.182***	0.137**	0.103	0.149**	0.093
	(0.057)	(0.065)	(0.077)	(0.061)	(0.097)
Observations	2974	2974	2974	2974	2974
Binary Y 2018	0.115***	0.087**	0.075*	0.082**	0.074*
	(0.031)	(0.038)	(0.044)	(0.035)	(0.041)
Observations	3014	3014	3014	3014	3014
Trimmed Sample 2018	0.656**	0.682**	0.683**	0.572*	0.470
(75 mi spillovers)	(0.329)	(0.330)	(0.330)	(0.340)	(0.372)
Observations	1351	1351	1351	1351	1351
Trimmed Sample 2018	0.688**	0.691**	0.693**	0.633*	0.711**
(100 mi spillovers)	(0.330)	(0.331)	(0.330)	(0.331)	(0.339)
Observations	987	987	987	987	987
Trimmed Sample 2018	0.259	0.211	0.204	0.228	0.165
(overlap)	(0.158)	(0.183)	(0.186)	(0.155)	(0.148)
Observations	683	683	683	683	683
Inverse hyper. Y 2020	0.518***	0.589***	0.467***	0.682***	0.301***
	(0.079)	(0.112)	(0.112)	(0.103)	(0.115)
Observations	2974	2974	2974	2974	2974
Binary Y 2020	-0.154***	-0.009	-0.020	0.019	-0.000
	(0.028)	(0.020)	(0.020)	(0.028)	(0.019)
Observations	3013	3013	3013	3013	3013
Trimmed Sample 2020	27.703***	23.465***	22.994***	25.885***	25.053***
(75 mi spillovers)	(6.214)	(6.474)	(6.553)	(6.271)	(8.299)
Observations	1427	1427	1427	1427	1427
Trimmed Sample 2020	29.194***	24.224***	25.130***	28.286***	28.477***
(100 mi spillovers)	(6.259)	(6.756)	(6.754)	(6.453)	(5.977)
Observations	1000	1000	1000	1000	1000
Trimmed Sample 2020	15.269**	15.336**	15.019**	21.135***	13.056*
(overlap)	(6.177)	(6.443)	(6.420)	(5.814)	(7.571)
Observations	674	674	674	674	674

Heteroskedasticity robust SE in parentheses; * p <0.1, ** p <0.05, *** p <0.01

Note: Robustness estimation results of the Average Treatment Effect on the Treated (ATET) from the context shock on protest events for the years 2018, 2019, 2020, and 2021. For the inverse hyperbolic Y specifications, Covariate Set 4 and for the binary Y specifications, Covariate Set 3 are applied. All other estimations are based on Covariate Set 1.

Robustness Salience Effect Year Variation

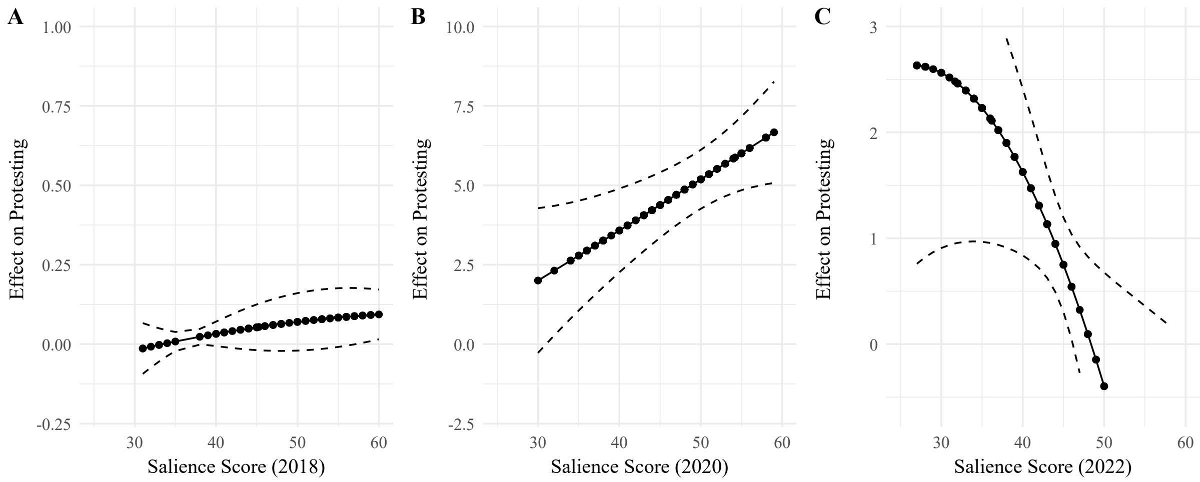


Figure 4.10: Estimation results fuzzy treatment of salience on protest event for the years 2018, 2020, and 2022. The dashed lines indicate the 95% confidence interval. LASSO selection Covariate Set 2 is applied.

Chapter 5

Conclusion

The collection of three research articles forms the basis of this dissertation. Each of them individually sheds light on a specific aspect of protest mobilization. In the first research article, different dimensions of resources are disentangled and assessed separately and in their interplay. The key conclusions are the identification of a substitutive effect between organizational capacities and online networks, as well as the mediating role of resources in the grievance effect. The second research article focuses on network effects and provides ample evidence for the importance of network considerations in the context of individual protest decisions. Finally, the third research article introduces a novel formal attention mechanism and presents empirical evidence for short-run attention-based protesting in the context of Black Lives Matter protests.

In combination, the three research articles present new theoretical and empirical insights about protest mobilization. While each article provides evidence on its own, they are connected as each of them investigates a potential source of selective incentives for protesting. None of them draws on holistic or deterministic explanation approaches for protesting. Instead, potential mechanisms that overcome the free-rider problem are introduced verbally or formally and empirically tested. In this sense, the dissertation presents novel insights into overcoming the free-rider problem and engaging in collective action in the form of protesting.

Bibliography

ABADIE, A. and CATTANEO, M. D. (2018). Econometric methods for program evaluation. *Annual Review of Economics*, **10** (1), 465–503.

ACEMOGLU, D. and ROBINSON, J. A. (2001). A Theory of Political Transitions. *American Economic Review*, **91** (4), 938–963.

— and — (2005). *Economic Origins of Dictatorship and Democracy*. Cambridge University Press.

AMERICAN NATIONAL ELECTION STUDIES (2021). ANES 2020 Time Series Study Full Release [Dataset]. www.electionstudies.org.

ANDREWS, K. T. and BIGGS, M. (2006). The dynamics of protest diffusion: Movement organizations, social networks, and news media in the 1960 sit-ins. *American sociological review*, **71** (5), 752–777.

ANDUIZA, E., CRISTANCHO, C. and SABUCEDO, J. M. (2014). Mobilization through online social networks: the political protest of the indignados in Spain. *Information, communication & society*, **17** (6), 750–764.

APOLTE, T. (2012). Why is there no revolution in North Korea? The political economy of revolution revisited. *Public Choice*, **150**, 561–578.

BAILEY, M., CAO, R., KUCHLER, T., STROEBEL, J. and WONG, A. (2018). Social Connectedness: Measurement, Determinants, and Effects. *Journal of Economic Perspectives*, **32** (3), 259–80.

—, GUPTA, A., HILLENBRAND, S., KUCHLER, T., RICHMOND, R. and STROEBEL, J. (2021). International trade and social connectedness. *Journal of International Economics*, **129**, 103418.

BARBERA, S. and JACKSON, M. O. (2019). A Model of Protests, Revolution, and Information. *Revolution, and Information (October 2019)*.

BBC (2018). Yellow vest protests 'economic catastrophe' for France. https://www.bbc.com/news/world-europe-46499996, last accessed: May 5, 2024.

BELLEFLAMME, P. and PEITZ, M. (2015). *Industrial Organization: Markets and Strategies*. Cambridge University Press, 2nd edn.

BELLONI, A., CHERNOZHUKOV, V. and HANSEN, C. (2014). Inference on treatment effects after selection among high-dimensional controls. *Review of Economic Studies*, **81** (2), 608–650.

BENNETT, W. L. and SEGERBERG, A. (2012). The logic of connective action: Digital media and the personalization of contentious politics. *Information, communication & society*, **15** (5), 739–768.

Beyer, J. L. (2014). *Expect us: Online communities and political mobilization*. Oxford University Press.

Bilgin, G. E. and Destan, C. G. (2024). *Voting Under Salience Bias and Strategic Extremism*. Tech. rep., University of Bonn and University of Mannheim, Germany.

Bivand, R., Millo, G. and Piras, G. (2021). A review of software for spatial econometrics in R. *Mathematics*, **9** (11), 1276.

Bonomi, G., Gennaioli, N. and Tabellini, G. (2021). Identity, beliefs, and political conflict. *The Quarterly Journal of Economics*, **136** (4), 2371–2411.

Bordalo, P., Gennaioli, N. and Shleifer, A. (2012). Salience theory of choice under risk. *The Quarterly journal of economics*, **127** (3), 1243–1285.

—, — and — (2013). Salience and consumer choice. *Journal of Political Economy*, **121** (5), 803–843.

—, — and — (2020a). Memory, attention, and choice. *The Quarterly journal of economics*, **135** (3), 1399–1442.

—, Tabellini, M. and Yang, D. Y. (2020b). *Issue salience and political stereotypes*. Tech. rep., National Bureau of Economic Research.

Boulding, C. E. (2010). NGOs and political participation in weak democracies: Subnational evidence on protest and voter turnout from Bolivia. *The Journal of Politics*, **72** (2), 456–468.

Brady, H. E., Verba, S. and Schlozman, K. L. (1995). Beyond ses: A resource model of political participation. *American political science review*, **89** (2), 271–294.

Buchanan, J. M. (2003). Politics without romance. *Policy*, **19** (3), 13–18.

Bursztyn, L., Cantoni, D., Yang, D. Y., Yuchtman, N. and Zhang, Y. J. (2021). Persistent Political Engagement: Social Interactions and the Dynamics of Protest Movements. *American Economic Review: Insights*, **3** (2), 233–250.

Cantoni, D., Heizlsperger, L.-J., Yang, D. Y., Yuchtman, N. and Zhang, Y. J. (2022). The fundamental determinants of protest participation: Evidence from Hong Kong's antiauthoritarian movement. *Journal of Public Economics*, **211**, 104667.

—, Yang, D. Y., Yuchtman, N. and Zhang, Y. J. (2019). Protests as strategic games: experimental evidence from Hong Kong's antiauthoritarian movement. *The Quarterly Journal of Economics*, **134** (2), 1021–1077.

Caren, N., Ghoshal, R. A. and Ribas, V. (2011). A social movement generation: Cohort and period trends in protest attendance and petition signing. *American Sociological Review*, **76** (1), 125–151.

Carter, E. B. and Carter, B. L. (2020). Focal Moments and Protests in Autocracies: How Pro-Democracy Anniversaries Shape Dissent in China. *Journal of Conflict Resolution*, **64** (10), 1796–1827.

Castelli Gattinara, P., Froio, C. and Pirro, A. L. (2022). Far-right protest mobilisation in Europe: Grievances, opportunities and resources. *European Journal of Political Research*, **61** (4), 1019–1041.

Castelnuovo, E. and Tran, T. D. (2017). Google it up! a Google Trends-based uncertainty index for the United States and Australia. *Economics Letters*, **161**, 149–153.

Chernozhukov, V., Hansen, C. and Spindler, M. (2016). hdm: High-Dimensional Metrics. *R Journal*, **8** (2), 185–199.

Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R., Gong, S., Gonzalez, F., Grondin, A., Jacob, M., Johnston, D., Koenen, M., Laguna-Muggenberg, E., Mudekereza, F., Rutter, T., Thor, N., Townsend, W., Zhang, R., Bailey, M., Barbera, P., Bhole, M. and Wernerfelt, N. (2022). Social Capital I: Measurement and Associations with Economic Mobility. *Nature*, **608** (7921), 108–121.

Choi, H. and Varian, H. (2012). Predicting the present with Google Trends. *Economic record*, **88**, 2–9.

Clarke, D. (2017). Estimating Difference-in-Differences in the Presence of Spillovers. Working paper.

Congress, U. (2018). The Geography of Social Capital in America. [Data set].

Crowd Counting Consortium (2021). "Crowd Counting Consortium" [Dataset]. crowdcounting.org.

Crump, R. K., Hotz, V. J., Imbens, G. W. and Mitnik, O. A. (2009). Dealing with limited overlap in estimation of average treatment effects. *Biometrika*, **96** (1), 187–199.

Dalton, R., Van Sickle, A. and Weldon, S. (2010). The individual–institutional nexus of protest behaviour. *British journal of political science*, **40** (1), 51–73.

De Mesquita, E. B. (2010). Regime Change and Revolutionary Entrepreneurs. *American Political Science Review*, **104** (3), 446–466.

Dorff, C., Adcox, G. and Konet, A. (2023). Data innovations on protests in the United States. *Journal of Peace Research*, **60** (1), 172–189.

Downs, A. (1957). An Economic Theory of Political Action in a Democracy. *Journal of Political Economy*, **65** (2), 135–150.

DRUKKER, D. M., EGGER, P. and PRUCHA, I. R. (2013). On Two-Step Estimation of a Spatial Autoregressive Model with Autoregressive Disturbances and Endogenous Regressors. *Econometric Reviews*, **32** (5-6), 686–733.

DURANTE, R., MASTROROCCO, N., MINALE, L. and SNYDER, J. M. (2025). Unpacking social capital. *The Economic Journal*, **135** (667), 773–807.

EARL, J. (2015). The future of social movement organizations: The waning dominance of SMOs online. *American Behavioral Scientist*, **59** (1), 35–52.

— (2018). Technology and social media. *The Wiley Blackwell companion to social movements*, pp. 289–305.

— and KIMPORT, K. (2011). *Digitally enabled social change: Activism in the internet age*. Mit Press.

—, MCKEE HURWITZ, H., MEJIA MESINAS, A., TOLAN, M. and ARLOTTI, A. (2013). This protest will be tweeted: Twitter and protest policing during the Pittsburgh G20. *Information, communication & society*, **16** (4), 459–478.

ECONOMIDES, N. (1996). The economics of networks. *International Journal of Industrial Organization*, **14** (6), 673–699.

EDWARDS, B. and GILLHAM, P. F. (2013). Resource mobilization theory. *The Blackwell encyclopedia of sociology*.

— and MCCARTHY, J. D. (2004). Resources and social movement mobilization. *The Blackwell companion to social movements*, pp. 116–152.

—, — and MATAIC, D. R. (2018). The resource context of social movements. *The Wiley Blackwell companion to social movements*, pp. 79–97.

EGOROV, G. and SONIN, K. (2024). The Political Economics of Non-democracy. *Journal of Economic Literature*, **62** (2), 594–636.

ELLIS, C. J. and FENDER, J. (2011). Information Cascades and Revolutionary Regime Transitions. *Economic Journal*, **121** (553), 763–792.

ENIKOLOPOV, R., MAKARIN, A. and PETROVA, M. (2020). Social Media and Protest Participation: Evidence from Russia. *Econometrica*, **88** (4), pp. 1479–1514.

FUCHS, C. (2006). The self-organization of social movements. *Systemic practice and action research*, **19**, 101–137.

GAMSON, W. A., FIREMAN, B. and RYTINA, S. (1982). *Encounters with Unjust Authority*. Chicago: Dorsey Press.

Gillham, P. F., Lindstedt, N. C., Edwards, B. and Johnson, E. W. (2019). The mobilizing effects of economic threats and resources on the formation of local Occupy Wall Street protest groups in 2011. *Sociological Perspectives*, **62** (4), 433–454.

González-Bailón, S., Borge-Holthoefer, J., Rivero, A. and Moreno, Y. (2011). The dynamics of protest recruitment through an online network. *Scientific reports*, **1** (1), 1–7.

Granovetter, M. (1978). Threshold models of collective behavior. *American journal of sociology*, **83** (6), 1420–1443.

Grasso, M. T. and Giugni, M. (2016). Protest participation and economic crisis: The conditioning role of political opportunities. *European Journal of Political Research*, **55** (4), 663–680.

Grimm, L. G. and Yarnold, P. R. (2000). *Reading and understanding MORE multivariate statistics.* American Psychological Association.

Gurr, T. (1968). A causal model of civil strife: a comparative analysis using new indices1. *American political science review*, **62** (4), 1104–1124.

Gurr, T. R. (1970). *Why men rebel.* Routledge.

Hirano, K. and Imbens, G. W. (2004). The propensity score with continuous treatment. *Applied Bayesian modeling and causal inference from incomplete-data perspectives, edited by Gelman A. Meng XL. West Sussex: Wiley InterScience*, pp. 73–84.

Huber, M. (2023). *Causal analysis: Impact evaluation and Causal Machine Learning with applications in R.* MIT Press.

Imai, K. and Ratkovic, M. (2014). Covariate balancing propensity score. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, **76** (1), 243–263.

Imbens, G. W. (2000). The role of the propensity score in estimating dose-response functions. *Biometrika*, **87** (3), 706–710.

James, G., Witten, D., Hastie, T. and Tibshirani, R. (2013). *An introduction to statistical learning: with applications in R*, vol. 103. Springer.

Jenkins, J. C. (1983). Resource mobilization theory and the study of social movements. *Annual review of sociology*, **9** (1), 527–553.

Jo, J.-I. and Choi, H. J. (2019). Enigmas of grievances about inequality: Effects of attitudes toward inequality and government redistribution on protest participation. *International Area Studies Review*, **22** (4), 348–368.

Jolliffe, I. T. and Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosophical transactions of the royal society A: Mathematical, Physical and Engineering Sciences*, **374** (2065), 20150202.

Justino, P. and Martorano, B. (2019). Redistributive preferences and protests in latin america. *Journal of Conflict Resolution*, **63** (9), 2128–2154.

Kabacoff, R. (2022). *R in action: data analysis and graphics with R and Tidyverse*. Simon and Schuster.

Kadivar, M. A., Ketchley, N., Sotoudeh-Sherbaf, A. and Barrie, C. (2025). Online calls for protest and offline mobilization in autocracies: evidence from the 2017 Dey Protests in Iran. *European Sociological Review*, **41** (1), 84–96.

Kann, C., Hashash, S., Steinert-Threlkeld, Z. and Alvarez, R. M. (2023). Collective identity in collective action: evidence from the 2020 summer BLM protests. *Frontiers in Political Science*, **5**, 1185633.

Kassambara, A. (2017). *Practical Guide to Cluster Analysis in R: Unsupervised Machine Learning*, vol. 1. Sthda.

Kaufman, L. and Rousseeuw, P. J. (2009). *Finding Groups in Data: An Introduction to Cluster Analysis*. John Wiley & Sons.

Kelejian, H. H. and Prucha, I. R. (1998). A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *The journal of real estate finance and economics*, **17**, 99–121.

— and — (1999). A generalized moments estimator for the autoregressive parameter in a spatial model. *International economic review*, **40** (2), 509–533.

— and — (2010). Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. *Journal of econometrics*, **157** (1), 53–67.

Ketchley, N. and Barrie, C. (2020). Fridays of revolution: focal days and mass protest in egypt and tunisia. *Political Research Quarterly*, **73** (2), 308–324.

Khondker, H. H. (2011). Role of the new media in the Arab Spring. *Globalizations*, **8** (5), 675–679.

Kim, H. W. and McCarthy, J. D. (2016). Socially organized sentiments: Exploring the link between religious density and protest mobilization, 1960–1995. *Social Science Research*, **60**, 199–211.

Kurer, T., Häusermann, S., Wüest, B. and Enggist, M. (2019). Economic grievances and political protest. *European Journal of Political Research*, **58** (3), 866–892.

Larson, J. M., Nagler, J., Ronen, J. and Tucker, J. A. (2019). Social networks and protest participation: Evidence from 130 million twitter users. *American Journal of Political Science*, **63** (3), 690–705.

Lecy, J. (2023). NCCS Legacy Core Series. [Data set].

LeSage, J. (2015). Spatial econometrics. In *Handbook of research methods and applications in economic geography*, Edward Elgar Publishing, pp. 23–40.

Little, A. T. (2016). Communication technology and protest. *The Journal of Politics*, **78** (1), 152–166.

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

— (2000). Collective action cascades: An informational rationale for the power in numbers. *Journal of Economic Surveys*, **14** (5), 655–684.

Malik, A. and Awadallah, B. (2013). The economics of the Arab Spring. *World Development*, **45**, 296–313.

Mapping Police Violence (2025). Mapping police violence database. Accessed: 2025-07-01.

McCabe, B. J. (2013). Are homeowners better citizens? Homeownership and community participation in the United States. *Social Forces*, **91** (3), 929–954.

McCarthy, J. D. and Zald, M. N. (1977). Resource mobilization and social movements: A partial theory. *American journal of sociology*, **82** (6), 1212–1241.

McPherson, M., Smith-Lovin, L. and Cook, J. M. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, **27** (1), 415–444.

Meirowitz, A. and Tucker, J. A. (2013). People power or a one-shot deal? a dynamic model of protest. *American Journal of Political Science*, **57** (2), 478–490.

Mellon, J. (2013). Where and when can we use Google Trends to measure issue salience? *PS: Political Science & Politics*, **46** (2), 280–290.

Miller, S., Roll, S., Brown, C. T., Brugger, L. and Grinstein-Weiss, M. (2023). Which side (of the balance sheet) are you on? household financial resources and participation in the 2020 protests. *J. Soc. & Soc. Welfare*, **50**, 56.

MIT Election Data and Science Lab (2018). County presidential election returns 2000-2020.

Moreno-Medina, J., Ouss, A., Bayer, P. and Ba, B. A. (2025). Officer-involved: The media language of police killings. *The Quarterly Journal of Economics*, **140** (2), 1525–1580.

Morrison, D. E. (1971). Some notes toward theorg on relative deprivation, social movements, and social change. *American behavioral scientist*, **14** (5), 675–690.

Mundt, M., Ross, K. and Burnett, C. M. (2018). Scaling social movements through social media: The case of Black Lives Matter. *Social media+ society*, **4** (4), 2056305118807911.

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

NCCS (2024). Core Files (Public Charities, Private Foundations, or Other 501(c) Organizations). Available from Urban Institute, National Center for Charitable Statistics.

Neyman, J. (1959). Optimal asymptotic tests of composite hypotheses. *Probability and statsitics*, pp. 213–234.

Olson, M. (1971). *The Logic of Collective Action: Public Goods and the Theory of Groups, with a new preface and appendix*, vol. 124. Harvard University Press.

Opp, K.-D. (2009). *Theories of political protest and social movements: A multidisciplinary introduction, critique, and synthesis*. Routledge.

Ortiz, I., Burke, S., Berrada, M. and Cortés, H. (2013). World protests 2006-2013. *Initiative for Policy Dialogue and Friedrich-Ebert-Stiftung New York Working Paper*, (2013).

Pilny, A. N., Atouba, Y. C. and Riles, J. M. (2014). How do SMOs create moral resources? The roles of media visibility, networks, activism, and political capacity. *Western Journal of Communication*, **78** (3), 358–377.

Popper, K. R. (1994). *The myth of the framework: In defence of science and rationality*. Psychology Press.

Putnam, R. D. (1995). Bowling alone. *Journal of Democracy*, pp. 65–78.

Quaranta, M. (2018). Nonviolent Protest in Europe: The Role of Macroeconomic Conditions in Party and Union Members' Participation. *Ps: political science & politics*, **51** (1), 79–83.

Raleigh, C., Linke, A., Hegre, H. and Karlsen, J. (2010). Introducing ACLED: An armed conflict location and event dataset. *Journal of peace research*, **47** (5), 651–660.

Reuters (2019). French 'Yellow Vests' protests cost 0.2 percentage points of growth - Le Maire. https://www.reuters.com/article/idUSKCN1QH17O/, last accessed: May 5, 2024.

Rohlfs, J. (1974). A Theory of Interdependent Demand for a Communications Service. *Bell Journal of Economics and Management Science*, **5** (1), 16–37.

Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, **70** (1), 41–55.

Rotolo, T., Wilson, J. and Hughes, M. E. (2010). Homeownership and volunteering: An alternative approach to studying social inequality and civic engagement 1. In *Sociological Forum*, Wiley Online Library, vol. 25, pp. 570–587.

Schelling, T. C. (1960). *The strategy of conflict*. Harvard University Press: Cambridge, Mass, 2nd edn.

Shadmehr, M. (2018). Protest puzzles: tullock's paradox, Hong Kong experiment, and the strength of weak states. *Quarterly Journal of Political Science, Forthcoming*.

— (2021). Protest Puzzles: Tullock's Paradox, Hong Kong Experiment, and the Strength of Weak States. *Quarterly Journal of Political Science*, **16** (3), 245–264.

Shafiq, M. N., Mason, J., Seybolt, T. and DeLuca, K. (2014). Are Student Protests in Arab States Caused by Economic and Political Grievances? Empirical Evidence from the 2006–07 Arab Barometer. *Peabody Journal of Education*, **89** (1), 141–158.

Shirky, C. (2008). *Here comes everybody: The power of organizing without organizations*. Penguin.

— (2011). The political power of social media: Technology, the public sphere, and political change. *Foreign affairs*, pp. 28–41.

Somma, N. M. (2010). How do voluntary organizations foster protest? The role of organizational involvement on individual protest participation. *The Sociological Quarterly*, **51** (3), 384–407.

Souza, P. C. (2024). The Empirics of Network Models. In *Oxford Research Encyclopedia of Economics and Finance*.

Steinert-Threlkeld, Z. C., Mocanu, D., Vespignani, A. and Fowler, J. (2015). Online social networks and offline protest. *EPJ Data Science*, **4** (1), 1–9.

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

Truex, R. (2019). Focal points, dissident calendars, and preemptive repression. *Journal of Conflict Resolution*, **63** (4), 1032–1052.

U.S. Census Bureau (2022). American community survey 5-year estimates. [Data set].

Vâlsan, C., Goschin, Z. and Druică, E. (2023). The measurement of social capital in America: a reassessment. *Social Indicators Research*, **165** (1), 135–161.

Van Dyke, N. and Amos, B. (2017). Social movement coalitions: Formation, longevity, and success. *Sociology Compass*, **11** (7), e12489.

Verba, S., Schlozman, K. L. and Brady, H. E. (1995). *Voice and equality: Civic voluntarism in American politics*. Harvard University Press.

Wang, D. J. and Soule, S. A. (2012). Social movement organizational collaboration: Networks of learning and the diffusion of protest tactics, 1960–1995. *American Journal of sociology*, **117** (6), 1674–1722.

Welzel, C. and Deutsch, F. (2012). Emancipative values and non-violent protest: The importance of 'ecological'effects. *British Journal of Political Science*, **42** (2), 465–479.

Wiedermann, M., Smith, E. K., Heitzig, J. and Donges, J. F. (2020). A network-based microfoundation of Granovetter's threshold model for social tipping. *Scientific reports*, **10** (1), 11202.

Williamson, V., Trump, K.-S. and Einstein, K. L. (2018). Black Lives Matter: Evidence that police-caused deaths predict protest activity. *Perspectives on Politics*, **16** (2), 400–415.

Wooldridge, J. M. (2007). Inverse probability weighted estimation for general missing data problems. *Journal of econometrics*, **141** (2), 1281–1301.

— (2010). *Econometric analysis of cross section and panel data*. MIT press.

— (2016). *Introductory Econometrics: A Modern Approach 6rd ed.* Cengage learning.

Yin, C.-C. (1998). Equilibria of collective action in different distributions of protest thresholds. *Public Choice*, **97** (4), 535–567.

Zald, M. N., McCarthy, J. D. *et al.* (1987). Social movements in an organizational society. *New Brunswick: Transaction Books.*

Zeitzoff, T. (2017). How social media is changing conflict. *Journal of Conflict Resolution*, **61** (9), 1970–1991.