Repression in the Digital Age: Communication Technology and the Politics of State Violence

Anita Rosemary Gohdes

Mannheim, 2014

Repression in the Digital Age: Communication Technology and the Politics of State Violence

Anita Rosemary Gohdes

Inauguraldissertation zur Erlangung des akademischen Grades einer Doktorin der Sozialwissenschaften,
Fakultät für Sozialwissenschaften,
Universität Mannheim

verfasst von
Anita Rosemary Gohdes

Mannheim, Oktober, 2014

First Examiner: Prof. Sabine C. Carey, PhD
Second Examiner: Prof. Nikolay Marinov, PhD
Third Examiner: Prof. Dr. Nils B. Weidmann
Dean: Prof. Dr. Michael Diehl
Date of Defense: 19 December 2014

Summary

The effect of the digital revolution on citizens' ability to voice dissatisfaction with their government and to coordinate dissent via social media has been the subject of much recent research. Optimistic accounts have so far failed to address the salient fact that the state maintains *de facto* control over the access to social media, which means that new digital technology also provides abusive governments with tools to repress challengers. This dissertation investigates how states' strategies of violent repression are informed by the use of these opportunities to control the internet.

I identify two main forms of control, which are the restriction or disruption of the internet on the one hand, and digital surveillance on the other hand. States face a trade-off: they can either restrict access to the internet and with it diminish opposition groups' capabilities, or they can permit the digital exchange of information and monitor it to their own advantage. I argue that the choice of internet control affects the *type* and *scale* of state-sanctioned violence used against perceived domestic threats. The choice of *digital surveillance* as a form of control is likely to be used in conjunction with targeted acts of localised violence against those identified as critical to the future success of opposition movements. The availability of highly specified intelligence on the intentions and location of opposition leaders enables states to use *targeted violence*.

Where states have chosen to respond to critical domestic threats in the form of *censorship*, they will also be more likely to visibly demonstrate their authority through a *heightened use of violent repression*. In addition, censorship severely limits the choices for violent action on the side of the government, by restricting the state's own access to the required intelligence for selecting precise targets. Consequently, during periods of censorship, state-sanctioned violence is likely to affect the domestic population *indiscriminately*.

I present a global analysis of the relationship between internet disruptions and the level of state-sanctioned violence, confirming that states who use network disruptions are also more likely to abuse the rights of their citizens. Evidence presented in case examples provides contextual understanding for the variety of different digital control tools which states have at their disposal.

The full implications of the theoretical argument are tested by moving to the sub-national level, and investigating the relationship between internet control and state violence, spatially and temporally in the Syrian conflict. I present a new integrated database on incidences of state killing in Syria, as well as disaggregated measures of network accessibility. First, I show that internet shutdowns occur in conjunction with significantly higher levels of state violence, most notably in areas where government forces are actively challenged by opposition groups. Second, I use supervised machine-learning to analyze over 60,000 records of state killings by the Syrian regime, and classify them to distinguish between targeted and untargeted acts of repression. I show that

higher levels of internet accessibility are consistently linked to an increase in targeted repression, whereas areas with little or no access to the internet witness more indiscriminate campaigns of violence. I conclude the dissertation by discussing the implications of the theoretical argument and the results, which have important ramifications for research and policy attempting to limit state abuse in the twenty-first century.

Acknowledgements

This dissertation would not have been written without the help and encouragement of many brilliant people. All remaining errors are my own, but all the clever ideas were borne out of intellectual exchange.

I would like to thank my dissertation committee, Sabine Carey, Nikolay Marinov, and Nils Weidmann: Nikolay Marinov encouraged me to think big, and build an overarching argument. Nils Weidmann has continuously provided critical feedback on my work over the past years, and was instrumental for getting one of the dissertation chapters into shape for publication at the Journal of Peace Research. Sabine Carey has been the very best possible supervisor, mentor, guidance, and supporter I could ever have wished for. Her ability to give spot-on feedback, ask all the right questions, and move from capturing the big picture to the smallest detail is absolutely incredible. Her consistent guidance helped me turn initial ideas into a completed dissertation project. I feel lucky to have her as a role model. Thank you.

My dear colleagues at the Human Rights Data Analysis Group are the main reason why I was motivated to pursue further academic research on why and how governments abuse their power. Patrick Ball has been a true inspiration for my work in so many ways, not least due to his unapologetic commitment to the scientific truth. Thank you for taking me under your wings and teaching me the joys of programming! Megan Price is a wonderful colleague and co-author; working with her has helped me become a better scholar. I thank Jule Krüger for all her support throughout the last years, and I am happy to call her my friend.

A heartfelt thanks goes to my colleagues at the University of Mannheim, in particular Mascha Rauschenbach and Adam Scharpf, who provided me with so much laughter, feedback, and wisdom, and let me tap into their knowledge in countless conversations.

I would like to thank a number of scholars who provided crucial feedback to parts of this dissertation: Michael Colaresi, Kathleen Gallagher Cunningham, Thomas Gschwend, Cullen Hendrix, Bethany Lacina, Will Lowe, Jessica Maves Braithwaite, David Siegel, Jakana Thomas, and Elisabeth Wood. A special thank you goes to Will Moore and Christian Davenport, not only for their excellent feedback, but because they have continuously provided support and advice on the ins and outs of academia.

My co-authors Todd Landman and Johanna Birnir taught me how the process of publishing works and their advice has made me a better scholar.

I dedicate this dissertation to my parents, Rolf and Catherine Gohdes, whose love and tireless support cannot be measured in words, and who provided outstanding linguistic advice. I would not have been able to complete this dissertation without them.

Finally, I thank Lukas Stötzer for being the very best person. You are my favourite.

Contents

Summary **v**

Acknowledgements **vi**

1 Introduction **1**
1.1 Motivation and Research Question 1
1.2 Main theoretical argument 3
1.3 Empirical approach 4
1.4 Plan of the Dissertation 5
1.5 Central contributions 6
1.5.1 Theoretical and conceptual contributions 6
1.5.2 Empirical and methodological contributions 7

2 Broadening the repressive toolkit in the digital age **9**
2.1 Introduction 9
2.2 State repression: what it is and how it works 10
2.3 Communication technology and contentious politics 13
2.3.1 Revolution technology? 14
2.3.2 Censorship in the digital age 16
2.3.3 Digital surveillance, or 'dataveillance' 18
2.4 Summary 19

3 Theoretical framework **21**
3.1 Introduction 21
3.2 The logic of internet censorship 22
3.3 The logic of internet surveillance 27
3.4 Summary: Network control and violent coercion as concerted strategy 32

4 Global evidence: internet control and coercion **33**
4.1 Introduction 33
4.2 Four case examples 34
4.3 Internet disruptions and coercion: a global analysis 38
4.3.1 Data and empirical strategy 38
4.3.2 Results 42
4.4 Summary 45

5 Integrated data on state killings in the Syrian Arab Republic **47**
5.1 Introduction 47
5.2 The problem of over-counting violence in event data 48
5.3 Data 49

5.3.1 Integrating multiple sources . . . 49
5.3.2 Record-linkage . . . 50
5.3.3 Descriptive comparisons . . . 52
5.4 Summary . . . 54

6 Accounting for the dark figure: unreported violence in event data 57
6.1 Introduction . . . 57
6.2 Information access and the challenge of identifying trends in event data . . . 57
6.3 Modelling the reporting of violence . . . 60
6.4 Simulations . . . 63
6.5 An example: control, contestation and violence in Syria . . . 67
6.6 Summary . . . 70

7 The military strategic value of national network disruptions 73
7.1 Introduction . . . 73
7.2 Outages and operational advantages . . . 74
7.2.1 Testable implications . . . 78
7.3 Data and empirical strategy . . . 80
7.4 Analysis I: Network outages and documented killings . . . 81
7.4.1 Descriptive evidence . . . 81
7.4.2 National-level evidence . . . 82
7.4.3 Regional evidence . . . 84
7.4.4 Placebo Tests . . . 86
7.5 Analysis II: Network outages and documentation patterns of violence . . . 87
7.5.1 Variation in reporting before and during disruptions . . . 87
7.6 Summary . . . 89

8 Information, connectivity, and strategic coercion 91
8.1 Introduction . . . 91
8.2 Information control and coercion . . . 92
8.3 Surveillance, censorship, and 'effective' repression . . . 94
8.3.1 Incentives for surveillance . . . 95
8.3.2 Incentives for censorship . . . 96
8.4 Data and empirical strategy . . . 96
8.4.1 Regional network accessibility in Syria . . . 96
8.4.2 Classifying targeted and untargeted repression . . . 97
8.4.3 Estimating reliable levels of targeted and untargeted killings 99
8.5 Results . . . 100
8.6 Summary . . . 102

9 Conclusion 105
9.1 Theoretical contribution . . . 106
9.1.1 Surveillance, censorship, and violence after the digital revolution . . . 106
9.1.2 Network control and military warfare . . . 106
9.2 Methodological contribution . . . 107
9.2.1 Integration and classification of high-quality data on government coercion . . . 107

9.2.2 Addressing bias in documented event-data 107
9.3 Policy implications 107
9.3.1 Government accountability 107
9.3.2 Security implications for citizens and activists 108

Bibliography **109**

Appendix **125**

List of Figures

4.1 Egyptian internet traffic during the shutdown, Jan 25-Feb 5, 2011 (A= 28 January, B= 2 February). 34
4.2 The ecosystem of Bahrain's "IP spy" attacks (Figure 2 by Marczak et al. (2014: 5)). 36
4.3 Major government-directed internet disruptions (full and partial disruptions), 1995-2010. 39
4.4 Government-directed internet disruptions, and human rights performance, 1995-2010. 42
4.5 Network disruptions and repression, global analysis. Point estimates and 95% confidence intervals. 45

5.1 Individual sources, and integrated data, over time, Syria, March 2011 - April 2014. 53
5.2 Density of reported monthly killings, Syria, March 2011 - April 2014. 53
5.3 Density of reported killings, by governorate, Syria, March 2011 - April 2014. 54

6.1 An example of violence and reported violence, over time. 59
6.2 Violence, and three sources of reporting. 60
6.3 Reporting density across three sources. 61
6.4 Simulated levels of violence in urban and rural regions. 64
6.5 Example of simulated reporting of violence in urban and rural locations. 65
6.6 Simulations: reported and estimated violence. 66
6.7 Weekly reported violence in Aleppo & Damascus, 2012. 68
6.8 Reported and predicted violence in Aleppo & Damascus, 2012. . 69

7.1 Mean difference in daily killings between days with and without internet, Syria, March 2011 - September 2013. 82
7.2 Violence and network disruptions, Hama 2011 & Rural Damascus 2012. 83
7.3 Difference of means tests for days with and without disruptions. 83
7.4 Expected change in daily killings, given network disruptions. . . 85
7.5 Time-shifted placebo treatment test. 87
7.6 Per cent of undocumented fatalities (of actual number) one week prior to, and during disruptions, by governorate. 88
7.7 Per cent of undocumented fatalities one week prior to, and during disruptions, measured for each disruption separately. 89

8.1 Internet (DSL, 3G, and 2G) accessibility by Syrian governorate, June 2013 - April 2014. 97
8.2 Assembling information on record details. 98
8.3 Summary of classified records. 100
8.4 Expected proportion of targeted killings, given internet accessibility. 101
8.5 Expected proportion of targeted killings, given internet accessibility. 102
8.6 First difference: change in percentage of targeted killings (no access to full connection). 102

1 Major government directed internet disruptions (full and partial disruptions). 126
2 Disrupted Google Traffic, June 2011. 127
3 Disrupted Google Traffic, Nov/Dec 2012. 127
4 Disrupted Google Traffic, May 2013. 127
5 Expected proportion of targeted killings, by Syrian governorate, using the upper bound of estimated killings. 128
6 Expected proportion of targeted killings, by Syrian governorate, using the lower bound of estimated killings. 128

List of Tables

3.1 Network control and implications for coercive strategy 31

4.1 Summary statistics, a global analysis of network disruptions and violence state coercion 42
4.2 Network disruptions and state repression, random effects models 44

5.1 Snapshot of the database (anonymised) 52

6.1 Example of individual and multiplicative capture histories 61
6.2 Predicted percentage changes of violence during battle 69

7.1 Expected effects for network disruptions and violence 79
7.2 Summary statistics, documented fatality counts 81
7.3 National-level time-series model: Disruptions and violence . . . 84
7.4 First difference model: Network disruptions and changes in violence 86

1 Introduction

1.1 Motivation and Research Question

A few weeks before the first mass protests ensued across Syria in March 2011, the regime led by President Bashar Al-Assad lifted a large number of bans on social networking platforms, including Facebook and Youtube. Up to that point, the Assad regime had controlled the most regulated media and telecommunications landscape in the Middle East (OpenNet Initiative, 2009). The removal of such restrictions was a step that other repressive states, including China, considered unthinkable. Suddenly Syrian citizens were free to digitally voice their anger and resentment towards a despotic regime with an appalling human rights record. From the state's point of view, this would appear to be a recipe for disaster. Why, after years of extreme censorship, would a deeply autocratic government suddenly permit unrestricted access to, and exchange of, information?

The ability to connect via large social network platforms has been celebrated by social scientists, policy makers, and human rights groups across the world as an empowering new way for ordinary citizens to collectively mobilise against repressive rulers (see e.g. Cohen, 2009; Diamond, 2010; Castells, 2012; Bennett and Segerberg, 2013). In fact, the US State Department perceived the role of social media platforms in the fight for democracy to be so crucial in 2009, that it officially requested Twitter to reschedule planned maintenance work in order to provide full accessibility to activists during Iran's post-election protests (Landler and Stelter, 2009). A former deputy adviser to the White House National Security Council went so far as to suggest that Twitter should receive a Nobel Peace Prize for being one of the crucial 'soft weapons of democracy', and a 'megaphone' for citizen movements across the world (Pfeifle, 2009).

Two years later, when civilian uprisings spread like wildfire across the Middle East and North Africa, social media was declared the principal tool of the protest movement, with journalists and researchers proclaiming that in the twenty-first century, 'the revolution will be tweeted' (Hounshell, 2011; Lotan et al., 2011; Rasha, 2011; Else, 2012). In consequence, the opportunities offered by the digital media to previously marginalised voices of dissent, and

the role they play in facilitating protest and resistance, have become the subject of extensive systematic research (see Garrett, 2006; Aday, Farrell and Lynch, 2010; González-Bailón et al., 2011; Lynch, 2011; Tufekci and Wilson, 2012; Howard and Hussain, 2013*a*). The fact that anyone with a working internet connection can now access, generate, and exchange content on the internet has been termed a real 'game changer' (Bellin, 2012: 127) for authoritarian regimes intent on maintaining control during mass popular protest.

Amidst all the euphoric accounts of the digital revolution, a crucial question remains unanswered: why power-hungry states, with de facto control over citizens' access to social media, should appear to impassively concede to defeat by these new tools. The simple answer is: they do not. Public attention has been focussed on the protest-enhancing elements of social media, but behind the scenes, governments across the world have been extremely active across the past two decades. They have been continuously developing and refining a whole arsenal of tools to surveil, manipulate, and censor the digital flow of information in the realm of their authority (see Deibert, 2008, 2010; Howard and Hussain, 2013*b*). *How these tools of digital control inform states' larger strategies of repression is the subject of this dissertation.*

This dissertation attempts to answer the question of how states' repressive strategies are affected by the digital revolution. I investigate how the opportunities and challenges of internet control affect the most widely-used tactic of repression, which is state-sanctioned violence. Governments intent on maintaining power over all adversaries, have long since combined the use of information control and restriction with the use of violence against those deemed threatening to their authority. Dictatorial regimes in the twentieth century, from Germany's *Third Reich* under Adolf Hitler to the military juntas waging Argentina's *Dirty War*, were masters in the art of influencing public opinion via the censoring and manipulation of mass media. The politics of media control are thus inadvertently linked to governments' concerted strategies of repression; restrictions on the media enable the use and justification of state violence (Davenport, 1995; Van Belle, 1997).

Digital communication technology has fundamentally altered the opportunities and benefits of information control, and the role it plays in strategies of state repression. As social media enthusiasts have rightly noted, the digital revolution has given individuals all over the world a platform to amplify their political claims and make themselves heard. The fact that states from Bahrain to Vietnam have invested enormous resources in censoring content and arresting digital opposition activists speaks volumes about the threat they perceived it to be. In September 2014, when protests erupted in central Hong Kong, Chinese authorities quickly blocked access to the picture-sharing service Instagram. In May 2014, Turkey's prime minister effected a ban on Twitter after it had become widely popular in the Gezi Park protests eight months earlier. State authorities from Egypt to Iran have shut down all internet services in response to internal unrest. At the same time, it would be naive to assume that repressive states would generally permit and support the expansion of the internet if they did not anticipate clear advantages for their own survival (see Rød and Weidmann, forthcoming). A former head-of-department in the Ministry for *Staatssicherheit* recently contended that the current United States National Security Agency's surveillance techniques would have been a 'dream come true'

for the secret service of the German Democratic Republic, where the number of wiretapped phones was restricted due to the country's limited technological capacity (Schofield, 2013). The digital age has made these surveillance tasks more affordable and precise.

In February 2011, after decades of controlling a highly censored media landscape, the Syrian regime realised that it had manoeuvred itself into an informational vacuum concerning the identity and extent of internal dissent simmering within its borders. Assad's clan was in dire need of a way to identify potential opposition, their location, and planned protest behaviour, in order to effectively eliminate threats to the status quo. Allowing the population to freely converse on Facebook and Twitter offered a low-cost and effective way to expand surveillance and gain a clearer picture of state enemies (see MacKinnon, 2012: 64). Across the next three years, it then went on to periodically shut down the country's entire access to the internet, at points where it deemed this necessary. In short, Syria's government decided to provide access to social networking sites, but it simultaneously increased its own intelligence for counterinsurgency operations, while also providing the opposition with new ways of organising of collective action.

1.2 Main theoretical argument

The digital age presents governments who fear for their political survival with a dilemma. On the one hand, dissidents and opposition groups are empowered through the use of social media; on the other hand these platforms offer themselves to previously unseen levels of surveillance and manipulation. States face a trade-off: they can either restrict access to the internet and with it diminish opposition groups' capabilities, or they can permit the digital exchange of information and monitor it to their own advantage. Both strategies of internet control - censorship and surveillance - cannot be implemented at the same time.[1]

This trade-off informs states' use of violence. I argue that the choice of internet control affects the *type* and *scale* of state-sanctioned violence used against perceived domestic threats. Choosing either censorship or surveillance as a form of digital control inevitably *limits* the use of some forms of violence and *enables* the use of other forms. Where states have chosen to visibly respond to critical domestic threats through a demonstration of control in the form of *censorship*, they will also be more likely to visibly demonstrate their authority through a *heightened use of violent repression*. Second, censorship severely limits the choices for violent action on the side of the government by censoring its own access to intelligence on precise targets. During periods of censorship, state-sanctioned violence is likely to affect the domestic population *indiscriminately*.

The choice of *digital surveillance* is likely to be used in conjunction with targeted acts of localised violence against those identified as critical to the future success of opposition movements. The availability of highly specified intelligence on the intentions and location of opposition leaders enables states to use *targeted violence*. Digital surveillance measures are likely to be linked to the use of targeted, individualised state-sanctioned violence.

[1] There are a few exceptions, such as methods used by the Chinese government. This anomaly is addressed in Chapter 4.

1.3 Empirical approach

Empirically, I expect state forces to employ targeted campaigns of coercion in areas where they grant citizens free access to information through the internet. Where connectivity is limited, I expect to observe less targeted, and higher intensity strategies of violence. I test these implications in a variety of ways, and present a number of solutions to empirical challenges related to the measurement of both internet control and state violence.

Measuring internet control and state violence is highly challenging, because both phenomena, by the nature of their subject, are not *intended* to be fully observable. Sometimes, states want their citizens to know they are controlling the internet, and sometimes they do not. States sometimes want their citizens to witness state violence, and sometimes not. To empirically test the relationship between internet control and state violence, I present systematic cross-national evidence over time, as well as two distinct quantitative sub-national case studies, and four short qualitative case examples.

The observable implication that can most reliably be measured at a global level is the implementation of restrictions on the internet, and the associated increase in state-sanctioned violence. I present evidence from a global analysis of the relationship between internet disruptions and the level of violent state repression, confirming the hypothesised positive relationship. Taking into account the most important determinants for violent state repression, the evidence suggests that the implementation of internet disruptions is systematically linked to higher levels in rights abuses. Contextual evidence presented in the individual case examples combines qualitative evidence on the link between surveillance and state violence; it also reveals the variety of different means for digital control which states have at their disposal.

The full implications of the theoretical argument are tested by moving to the sub-national level, and investigating the relationship between internet control and state violence, spatially and temporally. The case selected for this analysis is the Syrian Arab Republic (known as Syria). I present a new integrated database on incidences of state killing in Syria, as well as disaggregated measures of network accessibility. The first study investigates the logic of internet outages for state killings. Network shutdowns occur in conjunction with significantly higher levels of state violence, most notably in areas where government forces are actively fighting violent opposition groups. To eliminate competing explanations, I estimate the number of undocumented killings prior to and during outages, to test whether disruptions are implemented to hide atrocities from outside observers. I find no support for this hypothesis. In the second study I use supervised machine-learning to analyze over 60,000 records of state killings by the Syrian regime, and classify them according to their event circumstances, to distinguish between targeted and untargeted acts of repression. To account for reporting biases, I estimate the actual number of incidences in each category. The results reveal that higher levels of internet accessibility are consistently linked to an increase in targeted repression, whereas areas with little or no access to the internet witness more indiscriminate campaigns of violence.

Case Selection

Syria was selected as the main case study for this dissertation, as it has been termed 'the most socially mediated civil conflict in history' (Lynch, Freelon and Aday, 2014: 5), with events being painstakingly captured, documented and communicated via the internet. Thousands of Youtube videos record the images of killed and injured people in morgues, hospitals and market places, and activists within and outside the country use countless Twitter and Facebook accounts to inform each other about military operations and massacres, and to organize and coordinate the revolution (see Youmans and York, 2012). The Syrian government has made use of a multitude of internet controls: it has fully disrupted all access, it has regionally limited the accessibility, and it has used an array of spying software to surveil its entire population. The regime employs an *Electronic Army* to enforce the regime's message throughout the virtual world. Under the banner of the 'Syrian Presidency', President Bashar Al-Assad also maintains a lively Instagram account with no discernible sign of the ongoing war.[2]

Social media use in Syria and events surrounding the Arab Spring indicate that the central role of the internet will only increase in future conflicts. For this reason, understanding the motivation behind network manipulations instituted by regimes fearful of their political demise will become an indispensable tool for our theoretical and practical understanding of conflict dynamics. Learning from current cases such as Syria is an important place to start.

1.4 Plan of the Dissertation

Chapter 1 provides a brief discussion of the main theoretical and empirical contributions made by this dissertation. *Chapter 2* reviews existing approaches to the study of state repression and research on the role of the digital media and contentious politics. Research on the effects of the digital media revolution on state repression in the digital era is an area that has remained totally underdeveloped. The missing link between communication technology and state violence is particularly astonishing given the growing body of research on the occurrence and variation in the use of digital censorship (e.g. Deibert, 2010; Howard, 2010). The chapter discusses digital censorship and surveillance in detail, addressing possible reasons for the lack of research on the link between internet control and state violence.

Chapter 3 sets out a first theoretical framework linking the logic of state-implemented internet controls to the use of strategic, violent coercion. The costs and benefits of both censorship and surveillance are discussed and their implications for state-sanctioned are considered.

Chapter 4 offers four distinct case examples of countries that have made use of network restrictions and surveillance techniques, in conjunction with clamping down on groups and individuals deemed threatening to them. It then presents systematic global evidence on the relationship between state-implemented network disruptions and the level of violent state repression between 1995 and 2010.

[2] See `http://instagram.com/syrianpresidency`.

Chapter 5 presents a new database on state killings in the Syrian Arab Republic (between March 2011 and April 2014). It addresses one of the main challenges faced by those collecting event data on violence: the over-counting of incidences. The second main challenge, namely, the under-counting of incidences, is addressed in *Chapter 6*. It presents an estimation solution to fluctuating dark figure of unreported violence, by modelling the reporting process, and predicting what went unreported.

Chapter 7 investigates the military strategic advantages of internet disruptions, and makes use of the just presented data and estimation strategy. Governments have a strategic incentive to implement internet blackouts in conjunction with larger repressive operations against violent opposition forces. Short-term intermissions in communication channels are expected to decrease opposition groups' capabilities to successfully coordinate and implement attacks against the state, allowing regime forces to strengthen their position. A competing explanation is that states implement blackouts to commit atrocities that are hidden from international scrutiny. I compare the level of underreporting before and during disruptions and find no evidence for this alternative hypothesis. In addition, the analysis implements a series of placebo tests to rigorously test the actual causal effect of the argument.

Chapter 8 looks at the full continuum of possible internet controls and analyses the link between the type of network control and the type of coercion used. I use supervised machine-learning to classify state killings into targeted and untargeted acts of repression. Using this classified data, and the estimation strategy presented in Chapter 6, I find that higher levels of information accessibility are consistently linked to an increase in the proportion of targeted repression, whereas areas with little or no access witness more indiscriminate campaigns of violence.

Chapter 9 concludes the dissertation by highlighting the main findings of the different chapters and discussing the main implications for future research and policy.

1.5 Central contributions

This dissertation breaks new grounds in investigating how states digitally inform their strategies of violent repression through the use of internet control. By doing so it contributes to the literature on state repression and the literature on the role of the internet in contentious politics, in a number of important ways.

1.5.1 Theoretical and conceptual contributions

Linking digital network control to state violence

To the best of my knowledge, this dissertation is the first study to theoretically and empirically investigate the link between state-implemented internet control and state-sanctioned violent repression. With the growing role of the new digital media in protest and opposition movements across the world, answers to the question of how states digitally inform their strategies of violence provide a crucial contribution to our understanding of the contemporary and future dynamics of protest and conflict.

Disaggregating digital information control

In looking at information control, the literature on state repression has focused on the causes and effects of traditional media censorship. The role of digital information control has received little attention, and has tended to be equated with internet restrictions. The logic of state surveillance has remained altogether under-researched. In this dissertation, I make the crucial differentiation between digital network restrictions, leading to censorship, and network provision needed for digital surveillance. Together, censorship and surveillance form the integral parts of states' options for internet control. This dissertation contributes significantly to our understanding of how states make use of these tools, and what costs and benefits are associated with them.

1.5.2 Empirical and methodological contributions

Integrating high-quality event data from multiple sources

Event data on political violence has recently been subject to substantial, important criticism. Quantifying violence is extremely challenging, and two of the main problems researchers face are the over-counting and the under-counting of events. This dissertation presents a newly integrated database on state killings in the Syrian conflict; it also addresses the fundamental problem of over-counting through the process of record-linkage. The new data offer the most complete and high quality assembly of reported state killings for the current conflict. By design, the database allows researchers to trace each recorded victim of state killing back to its original source, which makes it possible to investigate the actual process of reporting.

Accounting for unreported incidences in event data

The second problem in the collection of event data on violence is the under-counting of incidences. Empirical analyses using incomplete, unrepresentative data on violence run the risk of modelling the process by which violence was reported, not the dynamics under which it was perpetrated. Research on the relationship between information control and violence is particularly susceptible to this problem, as changes in information availability can affect our knowledge of violence, leading to problems of endogeneity. The main research question of this dissertation could not be addressed without a means of accounting for this problem. A statistical solution for predicting the under-counted incidences is presented in this dissertation. It is one of the first studies to use corrected statistical predictions of violence in multivariate analyses.

Supervised machine-learning classification of repressive strategies

Research on state repression and political conflict has made significant advances in disaggregating strategies of violence, yet the overwhelming majority of theoretical concepts are translated into a measure of 'body counts'. High levels of killing are automatically equated with indiscriminate violence, while low numbers are assumed to be selective. This dissertation contributes a new measure of repressive strategies by combining information on the circumstances

of each killing recorded in the new database and using supervised machine-learning algorithms to classify each incident as either targeted or untargeted. The opportunities for machine-learning techniques to be used in the analysis of large amounts of data are truly immense. The classifications presented in this dissertation exemplify the potential they offer for improving the overall quality of measures of political violence.

2

Broadening the repressive toolkit in the digital age

2.1 Introduction

The toolbox of instruments that states can use to repress their citizens has broadened dramatically with the rise of digital media and communication technology. Governments now have the option of controlling whether and in what form citizens are able to connect online, as well as the ability to extensively surveil peer-to-peer communication. Consequently, these measures provide efficient new methods for governments to effectively repress their citizens (Howard and Hussain, 2013*b*).

However, so far there has been a lack of research on how the state's use of violence is affected by these changes – despite the fact that there has been ample research demonstrating how the dramatic increase in collective organisation via social media platforms has made states more susceptible to both internal protest and dissent (Garrett, 2006; Earl and Kimport, 2011; Bennett and Segerberg, 2013).

In this dissertation, I argue that the information age has vastly increased opportunities for governments to gather detailed information on which to base their strategies for violent repression in response to these new forms of resistance. Faced with increased organizational abilities by the opposition via online communication, states will be more likely to disrupt network access in conjunction with a broad campaign of violent coercion. Conversely, where states perceive the threat of internal dissent to be low, they will likely maintain access to the internet in order to gain information on crucial opposition figures, who can then be targeted individually.

Classic literature on the determinants and dynamics of state repression has principally been concerned with the relationship between varying forms of state violence and domestic dissent (e.g. Lichbach, 1987; Mason and Krane, 1989; Davenport, 2007*a*). On the other hand, research on protest has principally focused on the ways in which individuals overcome coordination and collective action problems in order to organise sustained and effective resistance (e.g.

Lichbach, 1995; Wood, 2003). For more than a decade, studies concerned with the representation of individuals' grievances and the organisation of protest have intensely discussed how digital communication technology has changed the nature of contemporary protest movements, and asked to what extent previous theoretical models need to be updated or amended, given the social and political potential of the internet (see e.g. Surowiecki, 2005; Shirky, 2008, 2011; Earl and Kimport, 2011; van Dijk, 2012). The effect of the new digital media on citizens' involvement in contentious politics has thus become a crucial part of theoretical models on protest, perhaps most famously reflected by what has been termed the 'logic of connective action' (Bennett and Segerberg, 2013: 19).

In contrast, research on state violence in the digital era remains scarce. The missing connection between digital communication technology and state violence is particularly astonishing given the growing body of research on the occurrence and variation in the use of digital censorship (Deibert, 2008, 2010; Howard, 2010; Howard, Agarwal and Hussain, 2011; Howard and Hussain, 2013*b*; King, Pan and Roberts, 2013). This chapter reviews the current state of research on state repression, and provides an overview of how digital communication technology has influenced the dynamics of contentious politics.

I begin with a discussion of the definition and scope of state repression in general, and provide an overview of research on the main predictors for state violence. This is followed by an analysis of the research on the role of the digital media in contentious politics, and I address possible reasons for the lack of existing research on the link between internet control and state violence.

Chapter 3 then introduces the main theoretical argument of the dissertation, namely how differing forms of internet control, through the use of both surveillance and censorship methods, are linked to different strategies of coercive state violence.

2.2 State repression: what it is and how it works

Definitions of state repression vary in scope (for a full discussion see, e.g. Davenport, 2007*a*; Earl, 2011), but in general, this can be defined as:

> '[the] use of physical sanctions against an individual or organization [...] for the purpose of imposing a cost on the target as well as deterring specific activities and/or beliefs perceived to be challenging to government personnel, practices or institutions' (Goldstein, 1978: xxvii, cited in Davenport, 2007*a*: 2).

State repression is generally understood to be the violation of one or more of the basic human rights written into the International Covenant on Civil and Political Rights (Nowak, 1993). The types of sanctions applied by states can broadly be divided into civil liberty rights and physical integrity rights (Earl, 2011; Escribà-Folch, 2013). Civil liberty rights, sometimes also referred to as first-amendement type rights in the US context (see Davenport, 2007*a*: 2), include basic human rights such as the right to free speech, freedom of association, freedom of the press, and freedom of assembly. Physical integrity rights are rights that are meant to ensure every person's bodily integrity, including freedom from torture, freedom from being imprisoned for holding opposing political

views, freedom from execution by the state, and from being disappeared by force (Davis and Ward, 1990). The violation of physical integrity rights is also referred to as violent state coercion, and the terms will be used interchangeably in this dissertation.

The relationship between civil liberty rights and state coercion

In contrast to the above relatively broad definition, the majority of those doing *empirical* research investigating the determinants and dynamics of state repression have defined it as a narrow set of physical integrity rights violations, such as state killings, torture, political imprisonment and disappearance (e.g. McCormick and Mitchell, 1997; Poe, Tate and Keith, 1999; Moore, 2000; Danneman and Ritter, 2014; Fariss and Schnakenberg, 2014). The reason for this is that physical integrity rights violations on the one hand, and civil liberty violations on the other hand follow different dynamics, and can therefore *not* be used as substitutable indicators for repression.

The scope of targets which governments aim at varies substantially across different forms of repression. States can repress all citizens indiscriminately, or they can target specific individuals or groups, and the effects of repression can be either direct or indirect. In general, the violation of physical integrity rights directly affects a selected target population, namely those individuals who are directly harmed. Thus, if a prominent state dissident is detained and killed, she is the direct target of coercion, even though a broader group of activists with whom she previously worked might also be indirectly affected. In a similar way, if hundreds of protesters are killed in a state-perpetrated massacre, the direct targets are the protesters killed, even if the indirect targets might constitute other citizens subsequently deterred from taking to the streets.

In the case of civil liberty rights violations, such as the denial of freedom of speech and freedom of the press, the situation is different. This form of repression generally affects an entire population, even if individual perceptions on the severity of censorship can vary. For example, if the media is prohibited from publishing anything that is critical of the government, then theoretically this constitutes the denial of every citizen's freedom of expression, even if not everyone will perceive this to be a problem (Van Belle, 1997). Civil liberty violations are therefore usually visible for a broad audience, as everyone is affected by them. In contrast, states have the potential to hide physical integrity rights violations from the broader population, in particular when they are only targeting specific individuals (Davis and Ward, 1990; Pion-Berlin and Lopez, 1991).

For this reason, research on the determinants of state violence has treated violations or provisions of civil liberty rights as a key explanatory factor for the level and nature of coercion used by the government (Cingranelli and Richards, 1999, 2010; Conrad and Moore, 2010; Frantz and Kendall-Taylor, 2014). Conrad and Moore (2010: 46) investigate the conditions under which governments that have used torture in the past, will refrain from doing so in the future, and argue that the willingness of governments to grant their citizens freedom of expression will act as an indicator for the presence of important checks and balances, that should help monitor whether, and to what extent, states are abusing their monopoly over the use of force (see also Davenport and Armstrong, 2004; De Mesquita et al., 2005).

Online platforms granting ordinary citizens the ability to voice their dissatisfaction with government actions have turned the internet into an integral part of civil liberty provision across the world (Howard and Hussain, 2013*a*). At the same time, increased communication opportunities between individuals has facilitated mobilization, and with it, increased the risk of domestic dissent (Pierskalla and Hollenbach, 2013).

State coercion and domestic dissent

It has been both theoretically and empirically established that the most important predictor for increased state coercion is the presence of domestic dissent, the most extreme form of this interaction between government and opposition manifesting itself in civil war (Lichbach, 1987; Francisco, 1995; Rasler, 1996; Poe, Tate and Keith, 1999; Poe, 2004; Carey, 2009).[1] Poe (2004) provides a comprehensive theoretical account of the rationalist decision-making model underlying government repression. His work builds on the 'Most-Starr Decisionmaking Model' (Poe, 2004: 17), referring to early work by Most and Starr (1989) dealing with the cost-benefit calculations governments perform when faced with internal dissent. The model identifies two key factors that governments – assumed to be unitary and utility-maximising actors – take into account when deciding whether the use of coercion is a rational policy option or not. The first factor is the government's perceived domestic strength (S), and the second factor is the perceived threat (T) of the domestic challenge to destabilise the political status quo. Poe (2004: 17) demonstrates that governments will have an incentive to become active when a) the perceived threat T is larger than the state's own strength S, or when b) the perceived threat is increasing relative to the state's perceived own strength between t-1 and t (see also Most and Starr, 1989: 126-28). Poe (2004: 19) stresses that violent coercion is but one option for governments to deal with a (perceived) unfavourable Strength-Threat inequality, and that other factors - such as the regime type of a country - affect the willingness and ability to actually use violence against domestic threats.

Directly related to the determinants of state coercion is the question whether and under what circumstances violent coercion has in fact proven to be effective at suppressing and deterring further protest or rebellion (Lichbach, 1987; Moore, 1998; Carey, 2009; Lyall, 2010; Siegel, 2011; Davenport, forthcoming). Siegel (2011) contends that violent 'repression acts to disrupt the mobilization dynamic [in individuals' social network pathways] by removing participants and cutting these pathways' (Siegel, 2011: 997). Whether the coercive disruption of collective mobilization processes is successful or not depends on the reaction of those who are affected by it (Schutte, 2014). Groups enduring or witnessing coercion will either become fearful of further reprisals and reduce their activities (Lyall, 2010), or they will be motivated to increase their anti-government activities and rally further supporters for their cause in the same process (Lichbach, 1995; Kalyvas, 2006; Downes, 2008). For coercion to effectively work, governments have a

[1] The list of studies confirming and qualifying this relationship is long, including: Mitchell and McCormick (1988); Mason and Krane (1989); Poe and Tate (1994); Davenport (1995); Moore (1998); Poe, Tate and Keith (1999); Moore (2000); Zanger (2000); Regan and Henderson (2002); Poe (2004); Davenport (2007*b*); Pierskalla (2009); Carey and Gibney (2010); Carey (2010); Conrad and Moore (2010); Earl (2011); Conrad (2011); Danneman and Ritter (2014); Escribà-Folch (2013), to name just a few.

vested interest in using sufficient coercion to deter further dissident actions, while at the same time not motivating previously uninterested (or ambiguous) citizens to solidarise with the opposition.

To counter perceived domestic threats without alienating impartial citizens, governments need to be able to distinguish between threatening dissidents and the remaining, non-threatening population (Kalyvas, 2006). Where individual threats are identified, state forces will attempt to target and 'remove' those individuals, while leaving the remaining population unscathed. Where, however, the majority of the domestic population has been identified as threatening, violent coercion is likely to be wide-spread and indiscriminate (see Valentino, Huth and Balch-Lindsay, 2004).

Regime type and state coercion

In addition to the presence of dissent, the type of regime in power has been found to be robustly associated with differing levels of state coercion. Governments that operate in the domain of strong democratic institutions are the least likely to perpetrate violence against their population (Poe, Tate and Keith, 1999; Davenport, 1999; Zanger, 2000; Davenport and Armstrong, 2004). The presence of democratic institutions does, however, not necessarily mean that no acts of state violence will be committed. De Mesquita et al. (2005) show that physical integrity rights only benefit from fully developed democratic structures, and that political participation is the most crucial aspect for the respect of these rights. Democratic states are also more likely to use coercive measures that are less visible than autocracies, as they have an invested interest in being perceived as human rights protectors (Rejali, 2011; Conrad and Moore, 2012).

The relationship between regime types and physical integrity rights violations is, moreover, not linear. What has been termed the 'more murder in the middle' (Fein, 1995: 170)-hypothesis purports that regimes that are neither fully democratic, nor fully autocratic are the most coercive states, as they face the highest degree of insecurity, or perceived threat (Regan and Henderson, 2002; Carey, 2009).

The list of factors influencing state repression is much longer, including a country's previous level of rights abuse (Sullivan, Loyle and Davenport, 2012), economic wealth and population size (Poe, Tate and Keith, 1999), international human rights norms (Hathaway, 2007), trade agreements (Hafner-Burton, Tsutsui and Meyer, 2008), naming and shaming campaigns (Murdie and Davis, 2012), inequality (Landman and Larizza, 2009), and state-imposed sanctions (Peksen, 2009), to name but a few. What still remains unclear is how digital communication technology has affected the basic decision-making of when, and how, states use coercion.

2.3 Communication technology and contentious politics

The expanding use of social media platforms to organize and internationally amplify domestic protest has spurred a vast new literature on the importance of new media for citizens to voice grievances and advance democratic processes from the bottom upwards (for an overview, see Castells, 2012). At the same time, governments are becoming increasingly sophisticated in controlling the

domestic internet infrastructure to their own advantage. While government-led censorship of the internet is starting to receive increased attention, the pervasiveness and effects of network surveillance on repression remain largely under-researched. More generally speaking, the current literature has failed to theoretically or empirically address the link between forms of state-implemented network control and violent state coercion.

2.3.1 Revolution technology?

Recent civilian uprisings from Burma to Iran, to the so-called 'Arab Spring' have inspired an exponential increase in research on the role and general significance of modern communication technology for non-state actors in contentious politics (see e.g. Rheingold, 2008; González-Bailón et al., 2011; Lotan et al., 2011; Dewey et al., 2012; Howard and Hussain, 2013*a*; Browning, 2013). Within this body of research, digital communication technology and new digital media are seldom clearly defined, but generally they include the usage of peer-to-peer online platforms, such as Facebook, Instagram, Youtube, and Twitter, and voice over internet provider services (VoIP, such as Skype), as well as platforms that are used to collect and aggregate information relevant to protesters (crowd-sourcing). Mobile phone technology, while pre-dating New Media, is generally included in this definition, since the appearances of mobile phones capable of accessing the internet and providing location-based services (known as smartphones) has somewhat blurred their distinction.

The majority of research forming our current understanding of social media and civil unrest is based on the analysis of individual cases. Questioning groups who took part in the Egyptian Tahir Square protest in 2011, Tufekci and Wilson (2012) find that the majority of participants had learned about the protests via digital peer-to-peer communication. Furthermore, they find that 48% of all participants had 'produced and disseminated video or pictures from political protest in the streets' (Tufekci and Wilson, 2012: 373) via social media. They conclude that organization via new media platforms, most notably Twitter and Facebook had a profound impact on the propensity of individuals to join protests. In a study on the recent Tunisian revolution, Breuer, Landman and Farquhar (forthcoming) identify a number of reasons for why social media acted as an important 'catalyst' for mobilizing protesters. The new media provided an opportunity for digitally active users to maintain a constant information supply when the traditional media was being censored by the government, and it also acted as an important platform for different protest groups to coordinate and structure their anti-government campaigns effectively. In addition, posting invitations and successes of the protest movement online attracted further supporters, and lastly, documenting the government's coercive response led to a consolidation of domestic and international sympathy for the protesters. The availability of social media platforms is thus assumed to be an essential tool for effectively overcoming collective action problems and coordinating sustained protests against the backdrop of government-induced information shortages (Breuer, Landman and Farquhar, forthcoming: 1).

Chowdhury (2008: 8) analyses the crucial role online platforms played in ensuring the dissemination of critical information during the Saffron revolution in Burma, where the Burmese government heavily censored conventional news outlets. He emphasises the importance of citizen journalists in countries where

the traditional media is heavily censored and controlled by the government. Upholding the flow of information is not only crucial for the domestic population, it also keeps the international community informed. He argues that:

> '[t]here were far fewer deaths in [Burma in] 2007 than there were in 1988. It is possible that the Internet saved the lives of many protestors, because the Junta feared even greater criticism from images of troops killing monks and civilians. The presence of the Internet in a dictatorial regime may save lives.' (Chowdhury, 2008: 14).

Howard and Hussain (2011) review the role of the digital media during the first uprisings of what has become known as the Arab Spring, and conclude that modern communication technology provided the means for localised grievances to gain a collective and structured momentum that helped mobilise hundreds of thousands of people. Furthermore, they conclude that opposition groups are dependent on the international recognition, and access to the internet has facilitated the task of reaching out to foreign governments and international NGOs considerably (Howard and Hussain, 2011: 44). The role of information technology thus moves beyond being an aid for revolutionary ignition, it is assumed to also help legitimise such processes in the long-run by providing platforms for opposition groups to position themselves favourably in front of the international community (Keck and Sikkink, 1998).

Moving from protests to armed rebellions, Pierskalla and Hollenbach (2013) study cross-sectional effects of mobile network availability on the propensity for regional violent events. Analyzing cell phone coverage across African countries they find that locations with better access to wireless phone networks display higher numbers of violent events.[2] Taking a closer look at the insurgent side of internal conflict, the authors argue that cohesive rebellious activities are a challenge to coordinate – especially when groups are secretly operating across different locations – and therefore strongly benefit from the availability of cheap communication tools (Pierskalla and Hollenbach, 2013: 210).

Evidence from the violent unrests following the 2007 elections in Kenya adds a further explanatory layer of understanding to this process. Goldstein and Rotich (2008) investigate the use of mobile phones in the aftermath of the elections and find that they were not only used to coordinate violent events, they were also used to incite ethnic hatred. All across Kenya, citizens received text messages on their mobile phones motivating them to stand up and voice their dissatisfaction with the way the elections had been conducted. Goldstein and Rotich (2008: 8) recall one of the messages, which read:

> "Fellow Kenyans, the Kikuyu's have stolen our children's future...we must deal with them in a way they understand...violence." (Goldstein and Rotich, 2008: 8).

The use of media platforms to incite ethnic hatred is nothing new, one of the most prominent examples being the role of radio announcements in the Rwandan genocide (Kellow and Steeves, 1998), but the degree of precision with which individuals can be targeted by messages sent directly to their mobile phones is,

[2]Shapiro and Weidmann (forthcoming) highlight a potential weakness: Violent events are equally more likely to be reported where communication is facilitated by mobile networks.

however, new. Goldstein and Rotich (2008) also discuss the critical role citizen journalists took on during the crisis, by providing uncensored information when the national newspapers ceased to report about events on the ground.

Shapiro and Weidmann (forthcoming) argue that from the position of governments, increased networking opportunities also increase the possibility of civilians unwittingly sharing knowledge about planned opposition attacks with state forces, thus throttling rebellious actions. Analyzing cell phone usage in the Iraqi conflict, they find that insurgent violence is significantly lower where increased mobile communication is available. The availability of digital communication technology thus affects the information flow of both government and non-government actors. It can help non-traditional citizen journalists to have their voice heard in larger contentious discourses, where state forces are censoring traditional media outlets. It can also assist governments in gathering information shared by ordinary citizens concerning the location and planned activities of armed non-state actors or terrorist organizations.

The recent popular uprisings in the Middle East and North Africa have generated enormous interest in the role being played by new digital media in organizing local protest and rebellion. However, very little is known about how the regimes facing such challenges to the status quo use these new technologies to their own advantage.

2.3.2 Censorship in the digital age

Censorship, generally defined as the prohibition or partial suppression of the freedom of expression and freedom of the media, forms an integral part of restrictive and repressive state behaviour (Van Belle, 1997; Shadmehr and Bernhardt, 2013; Lorentzen, 2014). The focus in this dissertation is on censorship - restrictions, disruptions, and filtering - of the internet, and with it the new digital media.

Repressive governments have a long history of using nuanced, proactive methods of censoring online content that is deemed hazardous to maintaining their status quo (see MacKinnon, 2012). Case evidence indicates that incumbent regimes try to limit the potential for collective organization via the internet by manipulating and censoring information. For example, during the 2009 uprising, the Iranian government allegedly disrupted internet access in the immediate aftermath of the elections (Aday, Farrell and Lynch, 2010: 20-21). Furthermore, SMS text-messaging was blocked during the entire election period (ibid.).

More recent examples of this process can be found in Libya and Egypt, where the internet was cut off in response to anti-government demonstrations in 2011 in both countries (Edmond, 2011). In September 2013, in the midst of anti-government protests which were sparked over fuel prices, Sudan responded by disconnecting its citizens from the internet (Madory, 2013), and the Central African Republic witnessed brief intermissions of all internet connections in the midst of ongoing violent clashes in December 2013. Burma's regime shut off all connections in response to the monks' protests in 2008, and China proceeded to take its Xinjiang province offline during ethnic riots in 2009 (MacKinnon, 2012: 51). The use of internet limitations and outages by Assad's regime in Syria adds yet another case to the list of governments busy tightening the screws on their control of digital networks.

State control of the internet is widespread and growing, and can broadly be divided into three categories, or generations of control (Deibert and Rohozinski, 2010). The first generation presents the most basic form of consistently blocking content, while the second generation is more dynamic, case-specific, and the third generation involves more subtle ways of warantless surveillance and normative campaigns against critical information (ibid.). These three generations of network control types are not mutually exclusive, and can all occur within the same country, albeit at different points in time. First-generation controls are generally understood as the constant, static form of blocking full or partial access to the internet (Deibert, 2008). Examples of countries known for this type of censorship are Uzbekistan, Turkmenistan, the United Arab Emirates, and Saudi Arabia. In technical terms, these persistent blocking efforts are implemented by 'blocking access to servers, domains, keywords, and IP addresses' (Deibert and Rohozinski, 2010: 22).

In contrast, second-generation controls lend themselves to a more dynamic management of internet control. Deibert and Rohozinski (2010: 24-25) contend that many states have taken to implementing legal standards as to when, and under what circumstances, content can be blocked in the domestic cyberspace. The means by which second-generation controls are implemented are frequently not that different from those used in the first generation, but the blocking now occurs under political pretense, such as national and technical security interests. However, these pretense controls tend to require a higher level of technical sophistication, as they are implemented dynamically , termed "'just-in-time" or event-based denial of selected content or services' by Deibert and Rohozinski (2010: 25). What this means is that most of contemporary online censorship springs into action when and where the state perceives itself to be under imminent threat, such as during protests, strikes, or in post-election periods.

Event-based disruptions in response to perceived threats bring with them a range of strategic advantages, which will be discussed in more detail in section 3.2. One basic advantage these second-generation disruptions have over the static first-generation disruptions, is that because they are shorter in terms of duration, they can more easily be passed off as technical errors, or at least provide governments with grounds for plausibly denying active involvement in the disruption (Deibert and Rohozinski, 2010: 25). Nevertheless, in the majority of cases, specialists have been able to uncover the link to the government (Deibert, 2008, 2009, 2010).

The number of countries using second-generation internet controls is growing (Deibert, 2009), and in general these appear to also be the ones most afraid of their control being challenged through online protest mobilization, although the evidence collected to date is neither systematic nor complete (Deibert and Rohozinski, 2010: 28-29). These findings are reinforced in recent research by Lorentzen (2014), who provides a formal model for the logic of strategic censorship in authoritarian regimes, and asserts that such governments will only be motivated to censor the news when domestic unrest is prevalent, whereas censoring can be relaxed in peaceful times (Lorentzen, 2014: 403).

The risk of virtual communication inciting political unrest is corroborated by the behaviour of other non-democratic governments, the most prominent being China. In a large-scale quantitative analysis of social media censorship, King,

Pan and Roberts (2013) find that censoring in China is only aimed at comments and posts that could motivate collective action or advance the coordination of protests. In contrast, content criticising the government or its policies is not censored as it is not deemed threatening to the status quo. The level of sophistication involved in capturing and removing these specific statements suggests that the Chinese government perceives action-inciting comments as an actual threat to the regime's stability.

Third-generation network controls diverge fundamentally from the idea of actively censoring information, and instead focus on 'counterinformation campaigns [as well as the] active use of surveillance and data mining' (Deibert and Rohozinski, 2010: 27). Analyzing the modes of control in Russia and the Commonwealth of Independent States, the afore-mentioned authors find that highly authoritarian states are more likely to use traditional content-blocking to censor their cyberspace, whereas more democratic countries opt for less intrusive, surveillance-based approaches. The surveillance-based approaches will be discussed in more detail in the next section on digital state surveillance.

An extreme form of digital government censorship is the practice of shutting down entire domestic internet and cell phone services for shorter periods of time (see Deibert, 2008). Since these short-term disruptions are generally intended to address a specific political or social issue considered threatening to the government, they can be understood as part of second-generation internet controls. Howard, Agarwal and Hussain (2011) document 556 network outages between 1995 and 2011, across the world, with more than half of them occurring in authoritarian regimes. Their analysis suggests democratic governments generally shutdown internet access in an attempt to combat child pornography or other forms of overtly sexual online content. Authoritarian regimes, on the other hand, most frequently disrupt their networks in response to perceived national security issues, such as social and political unrest, or under the banner of moral or religious integrity and modesty (Howard, Agarwal and Hussain, 2011: 225). Since the regime type of a country is not directly related to the level of violent abuse the state exercises against its citizens, it is not entirely clear how the implementation of internet disruptions is related to the use of violence.

2.3.3 Digital surveillance, or 'dataveillance'

With the rise of peer-to-peer communication, digital control of the internet has taken on an important additional dimension, which is the extensive use of state surveillance of online communication and content exchange (Deibert, 2010). State surveillance can be defined as:

> 'the focused, systematic and routine attention to personal details for purposes of influence, management, protection or direction. Surveillance directs its attention [...] to individuals [...] it is deliberate and depends on certain protocols and techniques. [...] digital devices only increase the capacities of surveillance or, sometimes, help foster particular kinds of [it]' (Lyon, 2009: 14-15).

The rise of digital technologies has profoundly changed the capacity of states to surveil individuals, groups, and now even entire populations. The changes in state surveillance have gone so far, that they have on occasion been termed 'dataveillance' (Lyon, 2009: 16), referring to the large amounts of data

that are generated and stored in the process of surveilling individuals and their interactions with each other. Governments now have the ability to digitally register and store extensive details about individuals, such as their location, their friends, colleagues, consumption histories, fingerprints, and more recently, even DNA and fingerprints (Lyon, 2009). Monitored communication via SMS, email and social media platforms furthermore allows governments to not only construct dynamic profiles of individuals, but also to build a dynamic network model of groups and their information exchange across spatial, temporal, and topical dimensions.

The growing prevalence of invasive, warantless digital surveillance prompted the United Nations General Assembly to adopt Resolution 68/167 on the right to privacy in the digital age (UN, 2013). In June 2014, the United Nations Human Rights Office (OHCHR) issued an extensive report on the human rights implications of domestic and extraterritorial surveillance (Office of the United Nations High Commissioner for Human Rights, 2014). The report reiterates the growing ability of governments and companies to digitally monitor citizens all over the world:

> 'The State now has a greater capability to conduct simultaneous, invasive, targeted and broad-scale surveillance than ever before' (Office of the United Nations High Commissioner for Human Rights, 2014: 3).

Although the topic of unlawful interceptions and digital state surveillance of individuals and groups has received increased attention at the level of international human rights law, the subject remains largely under-researched in the field of state repression and political violence. The opportunities for states to use (or abuse) the information they gain from digital surveillance to target individuals and groups they deem threatening to the political status quo are, however, more than evident. In fact, the information exchange observable via social media platforms allows governments to not only identify those deemed most threatening to their own political survival, it also offers them information about the friends, followers and fellow activists who are most likely to sympathise with the dissidents' actions and beliefs.

Given all the advantages of internet surveillance in identifying targets, it has at least one major drawback for governments. In order to successfully monitor its citizens, a state needs to provide a high degree of internet accessibility for online interaction to actually take place. This in turns opens up the opportunity for citizens to use the connective strengths of social media to their own advantage.

2.4 Summary

Existing approaches to explain the occurrence and nature of state coercion in light of domestic dissent have failed to integrate the use of digital control into their theoretical models. This is particularly astonishing given the vast increase in literature on the way in which the new digital media has altered the dynamics of contemporary protest movements.

This chapter introduced definitions of state repression, coercion, and the characteristics of internet control through censorship and internet control through surveillance. I reviewed the literature on the main explanations for when and

why states use coercion against their own citizens. The role of the digital media in affecting contentious politics was discussed, focussing on research that demonstrates how non-state actors have benefitted from increased coordination and collaboration opportunities. I then presented an overview of different forms of internet control currently used by governments, and suggested how they are linked to states' goals of consolidation and maintenance of political power.

The next chapter introduces the main theoretical framework of the dissertation, which investigates the logic of digital censorship, through internet restrictions vs. the use of digital surveillance, through internet provision. It concludes with a discussion of the implications for state-sanctioned violence resulting from both forms of internet control.

3
Theoretical framework

3.1 Introduction

In this chapter, I present a theoretical framework that links the logic of state-implemented internet controls to the use of strategic, violent coercion. The theoretical approach is built on the assumptions made by existing rationalist decision-making models concerning state behaviour (Lichbach, 1987; Most and Starr, 1989; Poe, 2004). Governments faced with a domestic threat that is perceived to be either greater to their own perceived strength, or growing in comparison to their own perceived strength, will have an incentive to take action in order to regain control and maintain the status quo of political authority. The two policy options available to governments that are under consideration here, are the use of internet controls and the use of violent state coercion. Both policy options can be used to affect a behavioural change in the individuals or groups perceived to be threatening. For example, states can cut all domestic internet access (which means censorship), or they can choose to maintain internet access and monitor a specific group of individuals assumed to be dissidents (which means surveillance). Likewise, they can opt for an escalation of violence, by attacking and killing protesters indiscriminately, or they can choose to single out specific individuals deemed threatening and eliminate them in a targeted manner. In responding to perceived domestic threats, the government's choice of coercive strategy is unavoidably linked to the choice of internet control strategy. As will be discussed in more detail later, the two are linked because the choice of internet control strategy (censorship vs surveillance) inevitably *limits* the use of some forms of coercion and *enables* the use of other forms.

I argue that the violent implications of state-led internet controls can only be fully investigated if we differentiate between different types of network control *and* different types of strategic coercion. In short, modern-day governments faced with digitally mobilized citizens are presented with a fundamental trade-off: they can either *censor* their citizens' access to digital channels of mobilization, by restricting internet access, or they can provide this access and use the information gleaned from *surveilling* these channels to their own advantage. Both strategies

provide attractive benefits, but also come at considerable cost. I argue that the implications for a state's coercive strategy will be dependent on this tradeoff. Where governments restrict network access, they will also be more likely to employ broader, more indiscriminate campaigns of violence against their own population. In contrast, maintaining network connections in order to digitally surveil citizens will more likely be used where states are interested in identifying specific, individual threats, and therefore incidences of highly targeted state terror will be more prevalent.

Unlike the more traditional notion of *restricting* content through censorship, digital surveillance is based on the idea of *gathering* critical content through free-flowing information exchange among individuals. This distinction is both conceptually and empirically important, and has received much less attention in the literature on state-implemented internet control than the traditional forms of content blocking and filtering. Conceptually, state censorship rests on the idea that certain forms of information need to be banned from the public due to their potential for inciting dangerous actions among the population. Digital surveillance, in contrast, pre-requires a certain level of network access so that individuals can exchange their ideas and plans online, which the state then monitors and analyses.

To give an example, assume that a group of activists are planning an anti-government demonstration. A censorship strategy would assume that the protest can only be successful if the activists are able to reach out to fellow citizens via online social media platforms, and motivate them to join the demonstrations. Shutting down accessibility to these platforms should – according to this logic – reduce the successfulness and size of such a demonstration. Alternatively, states might decide to make use of digital surveillance. For surveillance to work, the activists would *have* to reach out to their fellow citizens via social media in order for the state to monitor the potential number of protesters, and register each individual's name, location, level of motivation to participate, and – if possible – history of anti-government activities. In the following section, I discuss the logic of these two different strategies of network control in more detail, and provide arguments about their constraining and enabling effect on the choice of violent, coercive strategy.

3.2 The logic of internet censorship

Disrupting full or partial access to the internet is, in itself, a policy that is low-cost, and quick to implement. Temporary digital blackouts can be excused as technical failures, giving governments – at least for a short time – the possibility to plausibly deny active involvement.

The benefits of this low-cost policy option are manifold. First, cutting digital communication channels is likely to significantly complicate the exchange of information that is critical of the government, making it increasingly hard for individuals to assess the extent to which fellow citizens are also frustrated with the political status quo. Second, the sudden absence of previously employed social media platforms means the collective organisation of dissent must revert back to slower forms of communication, which can lead to significant delays and inefficiencies for protest movements. Additionally, short-term disruptions of what citizens have come to regard as an integral part of a state's information

infrastructure provide an unmistakeable signal of 'who is in charge' of state power and control. Where governments make no secret of their censoring activities, timely disruptions (and the subsequent lifting thereof) also act as a punishment for a population that has dared to challenge the status quo. Lastly, in situations where governments are faced with armed internal challengers, such as an insurgency, or a terrorist group, the unexpected interruption of internet access can stifle these groups' military capabilities, by cutting off their access to important geographical services, such as Google Earth (see e.g. Keating, 2013).

Inhibiting information exchange

A principal reason why states censor information exchange is their fear that individuals may realise that a sufficient number of their fellow citizens are equally unsatisfied with the status quo. This in turn could lead to a lowering of any inhibition in joining anti-government activities, eventually resulting in serious challenges to state authority. The 'informational cascade model' formalised by Lohmann (1994: 49) assumes these events to unfold across different stages:

> '(1) People take costly political action to express their dissatisfaction with the incumbent regime. (2) The public then takes informational cues from changes in the size of the protest movement over time. (3) The regime loses public support and collapses if the protest activities reveal it to be malign.' (Lohmann, 1994: 49).

States have a vested interest in making the display of these informational cues as costly as possible, in order to inhibit the occurrence of such cascades (see also Kuran, 1989), but the rise in information exchange across social media platforms has evidently dramatically lowered these costs for anyone connected to the internet (Edmond, 2011; González-Bailón et al., 2011; Little, 2014). Restricting the exchange of information to avoid widespread diffusion during the critical second stage identified by Lohmann, can be a rational and easily implementable strategy for governments.

Inhibiting collective organisation

Collective organisation and coordination, as the previously discussed studies show, have dramatically been facilitated through the help of the interactive '2.0' internet. Evidence by King, Pan and Roberts (2014) on digital censorship in China also demonstrates that the power of 'connective action' (Bennett and Segerberg, 2013) is something governments fear even more than the mere exchange of information.

Thus, shutting down the internet for short periods of time not only stifles the spread of anti-government information, it also prevents individuals from collectively organising themselves, and from maintaining order and discipline during concerted protest actions. In addition, this disorder can give governments a reason to violently intervene to 'restore' order. If the disruption of digital communication channels is unexpected, governments have the advantage of surprising their opponents who have to regroup and coordinate activity via channels not dependent on the internet.

Signalling power and immediate punishment

The denial of basic services, such as electricity, water, waste removal and policing, can be an effective way to demonstrate governments' position of power towards political challengers. In the digital era, internet access has joined the list of basic services now open to state manipulation.

Where political challengers are gaining momentum in voicing their concerns about the legitimacy and credibility of incumbent regimes, displays of power and control can be an effective way to remind the domestic population of who has the capability to deny them basic infrastructural needs. Shutting down the internet in light of oppositional threats can thus act as an impressive demonstration of power, and equally important, of willingness to use it if need be.

Restriction of access to basic services also serves as a form of punishment for the broader population for allowing and possibly even supporting the formation of an opposition group that threatens the government. As with declarations of states of emergency, such as the Emergency Law under which Egypt was ruled for almost half a decade prior to the ousting of Hosni Mubarak, the denial of basic infrastructure can be blamed on threats to national security and the necessity for preventing any further harm.

State-imposed internet shutdowns are timely reminders of the de facto power of ruling governments, and can send a clear signal to the domestic population that further opposition activities are seen as a threat and will, in return, be heavily punished.

Inhibiting opposition capabilities

Governments faced with armed internal rebellion have a particular incentive to cut internet connections, and that is the stifling of the opposition's military capability. Recent conflicts in Libya and Syria provide extensive footage of opposition fighters using online mapping services, such as Google Earth and Google Maps, to accurately locate military targets, and to calibrate weapons to effectively reach said targets (Miller, 2012; Brownstone, 2011; Keating, 2013). New developments in geographical location systems made for personal use on devices such as smartphones and tablets have revolutionized the capacity to locate and target regime forces, with a level of precision that was not available a decade ago.

Cutting these connections will provide an operational advantage for state military and pro-government militas over oftentimes less well-equipped and ill-trained insurgent groups.

The cost of censorship

After demonstrating their ability and willingness to shut down internet services, states might perceive themselves as being powerful. However, a population that relies on the internet for both personal and professional reasons might soon become increasingly skeptical of a government that cuts them off from these channels. Worse, the denial of these basic services might even make a government appear to be somewhat desperate. Evidence suggests that following the outages which Egypt faced in 2011, an increasing number of protesters took to the streets all across the country, in order to demonstrate against Mubarak's

regime (Hassanpour, 2013). When Turkey's government decided to merely block access to Twitter in April 2014, the response was a national and international rallying of protests against Prime Minister Erdogan (Tufekci, 2014). Obstructing, or even just partly restricting accessibility can actually provide incentives for a neutral population to participate in anti-government protest. Ethan Zuckerman describes this effect in his 'cute cat theory', which suggests that:

> '[i]nternet tools designed to let ordinary consumers publish non-political content are often useful for activists because they are difficult for governments to censor without censoring innocuous content; because censorship of inoffensive content can alert non-activist users to government censorship.' (Zuckerman, forthcoming: 2).

Overly ambitious network disruptions to prohibit the organisation and coordination of a few select dissidents, may quickly dispel the illusion of freedom for the majority of internet users, in turn reinforcing the attention given to activists and broadening their platform. The restriction on network access might in effect provide the final impetus for ordinary citizens to take to the streets and protest against repressive government policies.

A further important drawback to shutting down the information access for the opposition, is that it also affects the state's own ability to gather intelligence about the nature and characteristics of the protesters. To effectively target internal threats, governments are largely dependent on intelligence that they collect through tip-offs from their civilian supporters, and through the monitoring of the actions and declared intentions of dissidents. As soon as these dissidents are no longer able to communicate online, states automatically have a harder time monitoring the opposition's plans and location. Furthermore, the lack of network access prevents supporters from providing critical information to the government.

Network disruptions can quickly attract huge international interest, as happened during the recent banning of Instagram in Hongkong, and the short-term ban on Twitter in Turkey. The outcry across the world, in particular on social media, was enormous when the internet was shut down in Syria, Libya, and Egypt. If social media is prohibited with the aim of reducing collaboration and coordination among domestic elites, then shutting down the entire internet supply might have the reverse effect of rallying further anti-government supporters. Since the international elite are connected around the globe via social media, the sudden blackout of information from a certain country might even produce a boomerang effect, whereby elites in other countries pressure their own government to condemn the actions in the repressive state (see Keck and Sikkink, 1998).

Lastly, but no less important are the economic losses that are associated with internet outages. When the Mubarak regime shut down internet services for five days in 2011, the Egyptian economy lost an estimated $90 million worth of revenues (Howard, Agarwal and Hussain, 2011: 231). This figure only includes direct losses in revenues due to the absence of internet and phone services, it does not include the shutdown of general communication services, such as those generated by tourism cites, call centers and e-commerce, as well as potential losses on investment in the aftermath of the blackout (see OECD, 2011).

Implications for coercive strategy

While the technical implementation of disruptions provides a relatively cheap policy option, the repercussions of cutting off internet access can clearly prove to be dangerously costly for the government. There are two main arguments that suggest why governments using internet disruptions in light of domestic threats will implement them in conjunction with violent coercive strategies that are *larger in scale and indiscriminate in terms of whom they target*.

First, from a government's perspective, the benefits of disruptions – namely gaining organisational and operational advantages over the opposition – will only be worth the costs, in situations where it perceives the domestic political threat to be *large*. Where a critical mass of citizens has already taken to the street and collectively organised a substantial amount of internal resistance and support, disrupting the internet can act as an immediate attempt to limit further diffusion of the protests, and prevent them from turning into population-wide uprisings. Where only a small portion of a domestic population is seen as challenging the status quo, the disruption of the internet would likely only lead to increased attention for those waging the anti-government campaign, and the disruptive response by the state could even lead to further support and solidarity with the protest movement.

Clearly, restrictions to digital communication only make sense if the expected support and solidarity that might possibly be generated towards anti-government groups in light of disruptions is negligible, when compared to the damage it is expected to cause to the opposition's capabilities to organise. This will most likely occur where the opposition has already reached a critical size and requires reliable communication channels to maintain its strength and momentum. The decision to visibly respond to critical domestic threats via a demonstration of control in the form of internet restrictions, suggests that a government is resolved and willing to counterattack the opposition with a heavy hand. As research on the logic of violence in civil conflict has revealed, state violence is likely to be indiscriminate in terms of who it targets where the state perceives the majority of its population as a threat (Valentino, Huth and Balch-Lindsay, 2004; Kalyvas, 2006).

The second reason that suggests internet disruptions will be accompanied by larger, indiscriminate campaigns of coercion relates to the constraining effects of the outage itself. Where a government has opted for the use of internet disruptions to avert further spread of unrest, it has simultaneously limited its own access to crucial intelligence significantly. Not only are anti-government groups barred from organising online, state forces now also lack access to this information. In addition, loyal civilian supporters of the government are prohibited from sharing knowledge about developments on the ground with them. In short, states sabotage their own access to information on the identity and location of the most 'dangerous' dissidents. The use of violence will inadvertently become increasingly indiscriminate.

Where states perceive the threat from opposition movements - regardless of whether these are armed or not - to be *large*, disrupting internet accessibility presents a somewhat desperate measure to avert further damage. Once the choice for active network disruptions has been made, and the technical implementation is completed, the options for using violence will be severely limited: without up-to-date intelligence on the developments of anti-government

activists, the government has less opportunity and capability to locate and target those individuals deemed most threatening. Strategies of violence employed by the state are then more likely to affect the domestic population indiscriminately, than to target dissidents individually.

3.3 The logic of internet surveillance

The technical means available for states to digitally surveil their population are manifold. Specialised software that allows governments to remotely intercept all online traffic by individual users (or groups) is sold by a variety of different companies, many of which are based in Europe and North America. For example, the *Gamma International Group* is based in the United Kingdom and Germany, and is best known for selling a product called *FinFisher* that provides governments around the world with a full surveillance toolkit (Wagner and Guarnieri, 2014).[1] Some governments employ their own software writers (or 'hackers') to produce spying software, also known as so-called spyware. Spyware is software that is designed for the use of spying on individuals or organizations. Governments also make use of software which is referred to as 'trojans' or 'trojan horses', which is generally distributed in the form of files or website links that, when executed, grant remote (and usually hidden) access to the user's device and information (see, e.g. Galperin, Marquis-Boire and Scott-Railton, 2013; Marczak et al., 2014).

The rising prevalence of surveillance being used for political purposes was discussed in a recent report by the UN Office of the High Commissioner for Human Rights, which states that digital surveillance is frequently used to specifically target political dissidents and members of the opposition (Office of the United Nations High Commissioner for Human Rights, 2014: 3). To achieve full surveillance, some countries have even passed mandatory laws for all owners of personal computers to install special software intended to filter politically sensitive content (ibid.).

Digital surveillance of entire populations, in particular of those identified as potential threats, is a highly rational policy option for governments who fear for their political survival. The fact that a vast amount of communication and generated information has been relegated to the digital world has made this form of surveillance increasingly lucrative in terms of the quality and quantity of information governments are able to track and (ab)use. The data gleaned from tracking online conversations can help identify dissidents in an early and precise way, providing governments with an opportunity to 'scotch' dissidents' activities before they grow and diffuse. Where protest does erupt, surveilling the entire population's response to it can help anticipate the potential for future dissent

[1] Global Voices Online reports that a recent inquiry to the German parliament revealed that between 2003 and 2013, German surveillance software was officially sold to Albania, Argentina, Chile, India, Indonesia, Qatar, Kosovo, Kuwait, Lebanon, Malaysia, Morocco, Mexico, Norway, Oman, Pakistan, Russia, Saudi Arabia, Switzerland, Singapore, Taiwan, Turkey, Turkmenistan, the United States of America, and the United Arab Emirates (see Wagner and Guarnieri, 2014). A detailed overview of German surveillance technology exports, can be found in a recent report in *Der Spiegel:* `http://www.spiegel.de/politik/deutschland/deutsche-spaehtechnik-gabriels-ausfuhrkontrollen-bleiben-wirkungslos-a-987555.html`.

and assess how the protests are received at large. In contrast to the censorship methods previously discussed, surveillance requires internet accessibility.

Yet another aspect which states may consider: consistently ensuring access to a medium that is actively used for entertainment purposes, and that can be influenced by government ideology, is likely to foster a certain complacency about political issues in the broader population. Instances where states partially reveal their monitoring activities – in an attempt to signal their capacity – can also deter collective action by encouraging self-censorship.

Identifying dissidents

Knowing where the perceived threat is coming from, and who constitutes the most 'dangerous' members of it is a crucial component for governments who are resolved to stay in power. For this reason, autocratic regimes build extensive and powerful intelligence branches within their security services, such as the *Staatssicherheit* in the former German Democratic Republic, or the *KGB* in the former Soviet Union. As mentioned in the introduction to this dissertation, prior to the expansion of mobile communication technology and the internet, the surveillance capacities of governments, while frequently pervasive, were very costly as they required an intense usage of human resources. Only a limited number of wiretaps were technically feasible, and secret service agents were frequently relegated to physically eavesdropping on those identified as potentially threatening.

The digital revolution has vastly shrunk the costs in this field, which means that states can now follow the communication and production of online commentary by a far greater number of people than was previously possible. Individuals who are active in writing critical articles about the government and posting them online can now be identified almost immediately, and their interaction and connection with other activists recorded and analysed. Geographically locating individuals through tracing their IP addresses or the signals from their mobile phones has become a routine operation, and frequently leads to the arrest of online activists. State authorities across the world, including in countries such as Vietnam, Russia, Bangladesh, and Ethiopia, routinely harass and arrest bloggers for voicing critical opinions on government policies or conduct. In many cases, these arrests occur before any of these activists have even launched activities outside of the digital world (see Committee to Protect Journalists, 2014).

It is clear that in today's world, digital surveillance offers important and tempting opportunities for governments to gather timely and precise information on the identity of potentially threatening activists and dissidents.

Anticipating future dissent

The production of social media content has, to a certain extent, become a mirror of real-life, in that the number and nature of online commentaries tends to vary with actual events happening on the ground (Zeitzoff, 2011). Challenges to a state's political authority, for example through the display of riots or protests, will quickly become the subject of discussions in the online world. Using digital surveillance to closely monitor the attitudes of the broader population towards both the challenger's actions, and the government's initial reaction, offers an important barometer of a society's general disposition towards the contemporary

political climate. Use of the information exchanged on social media to build network models of interaction between current and potential dissidents has been a strategy used by governments in countries such as Bahrain, Syria, and Egypt, to name a few.

A principal problem for authoritarian regimes is that they tend to suppress mass opinion for so long that they end up without a clear understanding of the political attitudes held by the majority of their population. Lorentzen (2013) argues that consolidated authoritarian regimes actually have an incentive to permit local protest at fairly regular intervals in order to gather information about political dissatisfaction. In a simliar way, monitoring the broader public's reaction towards minor protests can significantly enhance the governments understanding of potential grievances simmering at the surface of the population; discontent might turn into a critical threat if not countered in a targeted way.

Opium for the masses

The arguments presented up to this point have largely been concerned with the political potential of the internet, but it should not be forgotten that the majority of users across the world use it as a source of entertainment (Zuckerman, forthcoming). Research on the consumption of Western media in the German Democratic Republic demonstrated that citizens in authoritarian regimes may, to a large extent, be perfectly content so long as they are permitted to use these 'controversial' media sources and modes of entertainment (Kern and Hainmueller, 2009), something Kern and Hainmueller (2009: 377) refer to as 'Opium for the masses'. In a recent study on the power of mass media in stifling the risk of civil war, Warren (2014: 112-113) argues that regimes intent on staying in power can use mass media platforms to advance their own political ideology, and hereby make use of a wide array of tools, such as images signalling powerful leaders, that should increase popular support among the broader population.

Although the internet can also be used as a tool to transport state-controlled media content, it is far from being as streamlined as traditional mass media. However, research suggests that a large proportion of autocratic regimes have been able to consolidate their power with the expansion of internet accessibility (Rød and Weidmann, forthcoming); an indication that political elites seem to have been successful in transmitting their political agenda via these channels. When faced with the trade-off between internet restrictions and surveillance through network provision, the importance of internet accessibility to foster complacency should not be under-estimated.

Signalling power

Lastly, states can partially reveal their surveillance capabilities to selected parts of the population in order to communicate their ability and power. When in January 2014 protesters on Kiev's central independence square *Maidan Nezalezhnosti* received a text message that they had been recorded as participants in a mass disturbance (Murphy, 2014), the government was sending a clear signal of power by revealing that they were monitoring their challengers.

In short, the partial revelation of surveillance techniques in and of itself becomes a repressive measure meant to threaten and deter further anti-government activity. The message sent implies that the government not only knows *who*

is challenging it, but that it has identified these challengers as a threat, and is willing and able to take action against them. The knowledge of surveillance and its consequences is intended to actively encourage self-censorship among the population.

The cost of surveillance

Digital surveillance is evidently a powerful tool that has the potential to counteract anti-government activity before it diffuses extensively, and at the same deter action by those who know they are being watched. Digital surveillance, however, comes at a high price: it cannot be effective where the internet is excessively restricted. Unrestricted access to the internet, in turn, provides opposition groups with all the advantages and potential so heavily discussed by Social Media enthusiasts (see Diamond and Plattner, 2012).

States are faced with a trade-off of either restricting and censoring the internet, or permitting access to it while monitoring all content creation and exchange for potential threats.

Implications for coercive strategy

Digital surveillance via the internet is powerful, but can heavily backfire in cases where the strategic advantages gained by opposition groups outweigh the usefulness of the information gleaned from monitoring them. Two arguments suggest why governments making use of digital surveillance will likely make use of *targeted, individualised strategies of violent coercion.*

First, monitoring and identifying dissent is useful when and where governments perceive a threat to be increasing without having reached a full critical momentum. Where the threat is not yet fully visible, surveillance data can play a critical role in identifying and locating those initiating dissenting action. Once opposition groups have become institutionalised with accepted leaders who speak freely and openly for the entire group, covert information becomes openly available; the added value of surveillance decreases in comparison to the benefits gained from disrupting network accessibility. In concerted efforts to counter growing domestic threats, digital surveillance is likely to be most effective when used in conjunction with targeted acts of localised violence against those identified as critical to the protesters' future success.

Second, during periods of intense surveillance, the collection of highly specified intelligence on the intentions and location of critical players in anti-government movements enables state violence to be more targeted and tailored towards individuals. Digital surveillance during full internet accessibility is therefore likely to increase states' use of targeted, individualised violence against domestic threats.

	Censorship ⇓ Network Restriction	**Surveillance** ⇓ Network Provision
Benefits for the government	▪ inhibit collective organization ▪ inhibit capability ▪ inhibit *information cascades* ▪ punishment (access denial)	▪ identify targets (precise & early) ▪ analyse networks of dissidents/protesters ▪ survey attitudes (potential threats) ▪ foster complacency
Potential costs	▪ depreciated credibility ▪ economic loss	▪ information exchange & collective action ▪ (international & domestic) audience costs
Technical means* [*non-exhaustive list]	▪ full network shutdown ▪ partial restrictions ▪ content filtering & blocking	▪ spyware (trojans & malware) ▪ remote control systems (interception) ▪ geographical (locating & tracking) ▪ network analysis, change detection
Implications for coercive strategy	⇓ **high-intensity, indiscriminate (untargeted) display of force**	⇓ **targeted, individualised use of force**

Table 3.1. Network control and implications for coercive strategy

3.4 Summary: Network control and violent coercion as concerted strategy

The choice between internet restrictions, leading to active censorship, and internet provision, which can be used for digital surveillance purposes presents states with a fundamental trade-off that ultimately constrains them in their choice of coercive strategy. This chapter has provided a first theoretical entry point to investigating the logic of either of these strategies of internet control, and has suggested a mechanism by which they are both linked to different types of strategies of violent coercion.

The costs and benefits of both censorship and surveillance that were discussed in detail above, are summarised in Table 3.1, to provide a brief overview.

The rising importance of the digital sphere for individuals and groups in organising protest against de facto state authority, has turned the internet into a dangerous tool for governments who are fearful for their own survival. It has, however, also turned the internet into a powerful source of control for the government itself. As explored in this chapter, strategies of internet control are increasingly informing the strategies of state coercion, and the importance of further research in this field will be critical to understanding repression in the future.

4

Global evidence: internet control and coercion

4.1 Introduction

The previous chapter established the theoretical underpinnings of this dissertation: states make use of a variety of different forms of network control, which are subsequently integrated into their larger repressive strategy. I argue that the choice between network restriction, with its censoring consequences, and network provision, with its opportunities for surveillance, affects the choice of coercive strategy used by governments to throttle perceived threats, ranging from protests, to opposition movements and armed insurgencies.

The current chapter reconciles two empirical approaches to establish preliminary support for this argument. Section 4.2 presents four brief case examples of government use of network control. Egypt, Bahrain, Ethiopia, and China were selected so as to represent a broad variety of government types and varying successes of securing political stability. Egypt's government under Mubarak was overthrown in the wake of protests that largely commenced via social media organisation. Bahrain's regime also faced protests, but managed to maintain its power, not least by using dissidents' online activities to collect information on their networks and contacts, which was then used to enforce further arrests. Ethiopia's level of internet penetration is among the lowest in the world, and yet the government has invested significant resources in controlling its network, and the people using it. Lastly, China's online infrastructure represents one of the most sophisticated and resource-intensive systems of closely-linked censorship and surveillance.

In order to establish the general scope and relevance of network controls in explaining state's coercive behaviour, Section 4.3 provides a systematic analysis of the relationship between internet shutdowns and physical integrity violations across 171 countries between 1995 and 2010.

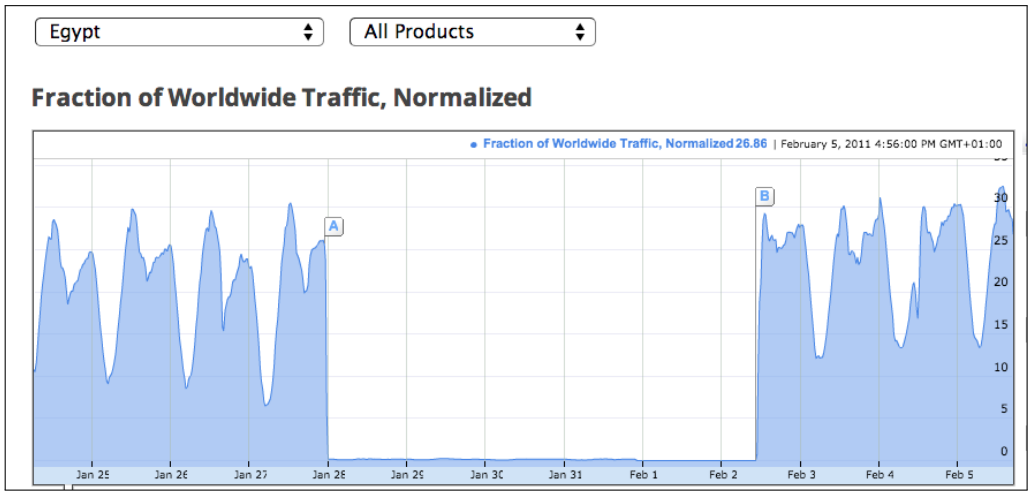


Figure 4.1. Egyptian internet traffic during the shutdown, Jan 25-Feb 5, 2011 (A= 28 January, B= 2 February).

4.2 Four case examples

Egypt

The ousting of Egypt's former president Hosni Mubarak in February 2011 has generally been attributed to the civilian uprisings that commenced in Cairo in January of the same year (Lynch, 2011). Social Media platforms such as Facebook, Twitter, and Youtube were heavily frequented by activists and citizens to facilitate organizing and coordinating protest (Hounshell, 2011; Lotan et al., 2011; Tufekci and Wilson, 2012).

The Egyptian government's response to this new 'threat' changed from surveillance to full censorship. To surveil activists, it made use of so-called *deep packet inspection (DPI) technology* (Singel, 2011) provided by the US-based Narus company, which allowed it to monitor all internet traffic throughout the protests in 2011, including Skype conversations, browsing histories and electronic messages (ibid.). Digital surveillance was conducted by the Egyptian State Security Investigation's Emergency Unit. In a report for Global Voice Online, Ramy Raoof writes that the Emergency Unit was regularly in charge of shutting down internet, mobile phone, and sms services in different cities or regions, blocking access to selected content, monitoring the activities of digitally active citizens, and of ensuring the swift cooperation of telecommunication providers (Raoof, 2011: online content).

Details of protesters' planned activities and locations gleaned from these data were repeatedly used to arrest and imprison opposition activists. The most prominently covered case was that of online activist and Google executive Wael Ghonim, who was arrested and secretly detained by the Egyptian authorities in January 2011 after being identified as the anonymous administrator of the 'We Are All Khaled Said' Facebook Page (Youmans and York, 2012: 318). Ghonim's Facebook Page had been one of the central communication platforms during the early days of the anti-regime protests in 2010.

On January 28 the regime changed its tactic of surveillance and implemented a country-wide shutdown of all internet connections. The blackout was imme-

diately picked up by a number of different services, including Google Traffic. Figure 4.1[1] shows the absence of traffic during this five day period. Connections were supposedly cut in an attempt to quell the spread of growing protests, and campaigns of violence against anyone participating in the unrest were intensified. The strategy quite obviously backfired and destabilized Mubarak even further, until he was forced to resign from office (Hassanpour, 2013). While the protests effected a change of power, the new government continues to arrest and detain online activists deemed threatening to the new political order (Mackey, 2014).

Bahrain

Despite initial large-scale protests in Bahrain's capital Manama in 2011 that were inspired by neighbouring revolutions, the Bahraini Al-Khalifa regime has managed to maintain its political power. Censorship of the media, as well as surveillance and arrests of dissidents and human rights activists form a central part of the Regime's repressive strategy (HRW, 2014*a*). Recent research by Marczak et al. (2014) identifies two types of attacks used against activists suspected of being involved in dissenting behaviour:

> 'The first involved malicious e-mails containing *FinSpy*, a "lawful intercept" trojan sold exclusively to governments. The second involved specially crafted *IP spy* links and e-mails designed to reveal IP addresses of operators of pseudonymous accounts. Some individuals who apparently clicked on these links were later arrested [...]' (Marczak et al., 2014: 3).

The authors were able to identify networks of attacks carried out by the government, where the Facebook user credentials from arrested activists were adopted to access and contact affiliated journalists and campaigners, who were then sent trojans containing spyware to facilitate targeted surveillance (Marczak et al., 2014: 5). The Bahraini regime has thus instrumentalised dissidents' online support networks as a referral-based information repository to target and arrest related activists. Figure 4.2 presents the network of surveillance and arrests identified by Marczak et al. (2014), which exemplifies how such attacks directly led to arrests of and house raids on government critics.

Ethiopia

Although less than one per cent of more than 90 million Ethiopians have access to the internet, the government has put in place a sophisticated system to control all online traffic. In 2011, it created the Information Network Security Agency (INSA), a part of the national security infrastructure that is steadily gaining discretionary power (HRW, 2014*b*: 29). *Reporters Without Borders* has accused INSA of being an 'NSA copycat'[2], referring to the agency's goal of accessing and monitoring all information Ethiopians exchange in the internet. To do this INSA has made use of an array of different spying software, including Remote Control

[1] Google Traffic Transparency Report: `https://www.google.com/transparencyreport/traffic/disruptions/30/`

[2] see `http://12mars.rsf.org/2014-en/2014/03/06/ethiopia-full-online-powers/`

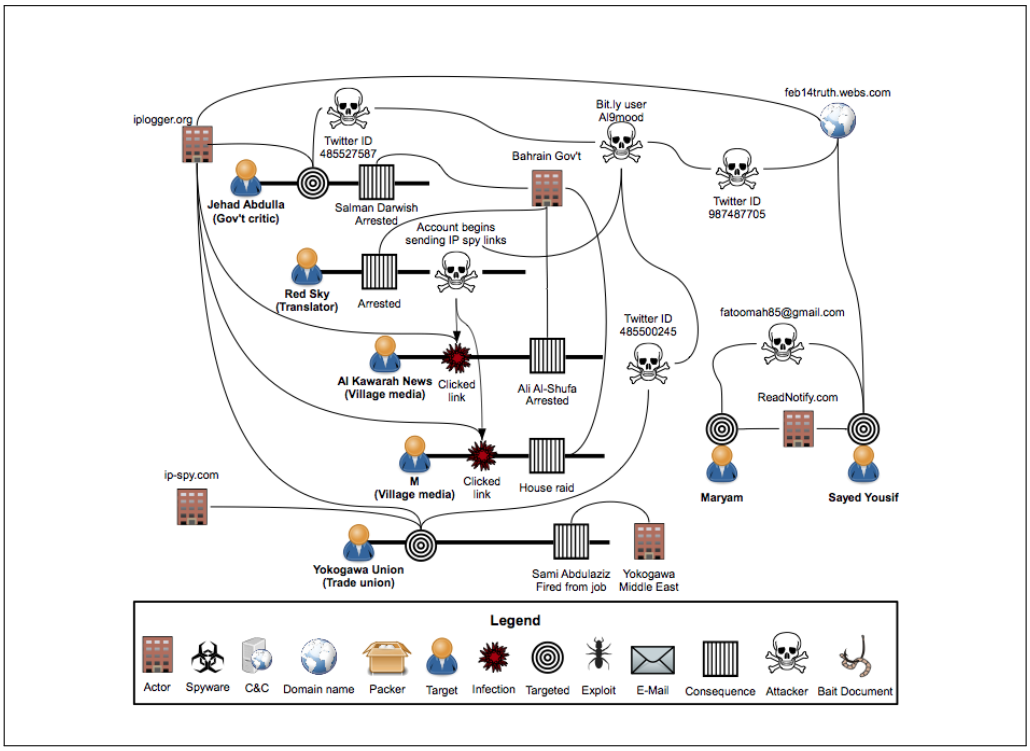


Figure 4.2. The ecosystem of Bahrain's "IP spy" attacks (Figure 2 by Marczak et al. (2014: 5)).

Systems (RCS)[3] and *FinSpy* interception software also used by the Bahraini government.

Parallel to its attempted overall surveillance, the Ethiopian government has repeatedly shut down cell phone and internet access in different regions of the country. As the state continues to exert full control over the country's telecommunication sector, selective disruptions or restrictions in anticipation of unrests can easily be implemented. In July 2013, Human Rights Watch interviewed a former Ethiopian government employee who reported:

> 'Whenever a demonstration is planned, the telecom service in eastern Harerghe is cut. During local elections it was cut. During recent Muslim protests it was cut. It is usually cut from 6 a.m. until after 2 p.m. Message I would get in Amharic is "for time being there is no service." Our network comes and goes all the time, but as soon as there is a problem for government there is no service whatsoever.' (HRW, 2014*b*: 50).

Human Rights Watch further reports incidences where individuals were geographically located through their cell phones and arrested for participating in protests. The location is frequently identified by locating the closest mobile tower utilised by the individual's phone to place the last call. Security officials are also known to harass suspected protesters by calling them multiple, successive times on their phones to determine which phone tower is receiving quick, successive phone calls (HRW, 2014*b*: 52).

[3] `http://www.hackingteam.it/index.php/remote-control-system`

It is clear that such locating strategies only work where the phone network has not previously been disrupted by the government, which means that potential protesters can either be located and surveilled, or prevented from communicating with each other, but not both at the same time. The Ethiopian government has evidently alternated between these two strategies on a frequent basis.

China

The People's Republic of China hosts one of the most sophisticated and all-encompassing systems for content-filtering, surveillance, and censorship in the world (OpenNet Initiative, 2012). Google, YouTube, Facebook, Twitter, and other major websites are blocked, but this does not mean that Chinese citizens do not make use of Social Media. On the contrary, the country boasts a large number of domestic networking platforms that are used by the majority of the online population, including Sina Weibo (offering services similar to Twitter), and Youku (offering services simliar to YouTube).

The scale of resources that has gone into building a domestic cyber-infrastructure makes China somewhat 'irregular' in its use of strategic censorship and surveillance. While most governments, at least to a certain degree, face a trade these-off between surveilling and censoring their population, the Chinese government has built its own infrastructure that allows for surveillance and subsequent censorship.[4] Research by King, Pan and Roberts (2013, 2014) demonstrates nevertheless that the Chinese government follows a highly strategic and careful logic of only censoring content that is intended to incite collective organization or action among the Chinese population. Content that directly criticises policies or state activities is, in contrast, censored at a much lower rate. The Chinese government's careful selection of content to be filtered and its reluctance to filter criticism against itself (where it involved no 'calls to action'), offers further support for the potential drawbacks of censoring content discussed in Chapter 3. Full censorship of information concerning the government, in particular in situations where the population is expecting critical discussions of certain events or policies, is likely to backfire and lead to a loss of credibility on the side of the state (see also Shadmehr and Bernhardt, 2013). Granting citizens a platform to exchange non-threatening criticism principally supports an illusion of free expression, strengthening the credibility of the government. The level of criticism posted online also acts as a barometer for the general level of (dis)satisfaction in the broader population.

On occasion, the government has, however, taken to more extreme measures than mere content censoring. In 2009, it proceeded to take its Xinjiang province offline during ethnic riots (MacKinnon, 2012: 51). Government officials contented:

> '[w]e cut Internet connection in some areas of Urumqi in order to quench the riot quickly and prevent violence from spreading to other places.' (Official in *Xinhua News*, quoted in OpenNet Initiative, 2012: 274).

[4] The simultaneous use of surveillance and censorship by the Chinese authorities is also likely to increase self-censorship (see Roberts, 2014).

Others contended that the government had cut the information flow to prevent reports of widespread arrests and interrogations (OpenNet Initiative, 2012: 274).

The case evidence set out above offers preliminary support for the arguments put forward in this dissertation: First, governments face a trade-off between digital censorship and surveillance. Second, digital censorship and surveillance are integrated into larger strategies of repression, and thus affect the nature of violent state coercion. The types of internet controls used vary widely across the cases, and yet all share a certain trait: surveillance methods and disruptions are linked to the use of violent coercion.

In the next section, I focus on the one observable proxy of internet control that can most reliably be measured at a global level: the implementation of restrictions on the internet. To establish the external validity of the main argument set out in this dissertation, I present evidence from a global analysis of the relationship between internet disruptions and the level of state-sanctioned violence.

4.3 Internet disruptions and coercion: a global analysis

Are governments that frequently disrupt domestic network services also more likely to use state-sanctioned violence against their own citizens? The evidence presented in this section suggests that in years where governments implement full or partial network disruptions, they are also significantly more likely to violate their citizens' physical integrity rights than in years where they permit uninterrupted internet access. The relationship is significant, even when taking into account the most commonly accepted confounders influencing state respect for human rights.

Internet shutdowns present a visible measure of de facto government network interference, and therefore open themselves to systematic cross-national analysis. In contrast, surveillance efforts during times of normal network provision can be conducted in secrecy, making global comparison a challenging task. The above-mentioned four case examples demonstrate the use of such surveillance technology in conjunction with state coercion, but since the majority of spyware and interception tools are, by definition, intended to work in secrecy, any quantitative cross-national comparison would be seriously flawed. Consequently, the following analysis aims at establishing empirical support for theoretical argument that network disruptions are likely to occur in conjunction with larger, indiscriminate campaigns of coercive violence exercised by the government against its population, as presented in detail in Chapter 3:

- **Empirical Expectation:** All else equal, governments that implement network disruptions are likely to abuse citizens' physical integrity rights at a higher level than governments that do not implement network disruptions.

4.3.1 Data and empirical strategy

The global prevalence of internet shutdowns

To assess the global prevalence of digital network disruptions, I construct different indicators that are based on an event-dataset collected by Howard,

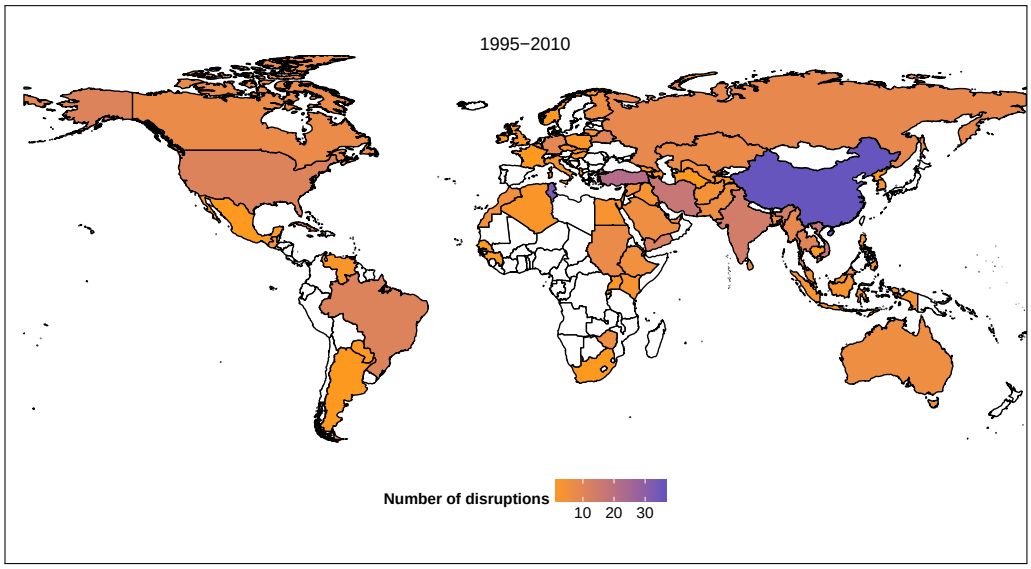


Figure 4.3. Major government-directed internet disruptions (full and partial disruptions), 1995-2010.

Agarwal and Hussain (2011).[5] Building on an analysis of news sources including international and domestic news sources, as well as expert sources such as information security blogs and specialised internet fora (Howard, Agarwal and Hussain, 2011: 221), a catalogue of all major disruptions and shutdowns of national digital networks was constructed. An event is defined as an instance of government-led disruption of the internet, leading to the shutting down of connections between the domestic and the international cyberspace, thus fully or partially disconnecting domestic network users (Howard, Agarwal and Hussain, 2011: 221).

The dataset distinguishes between four types of disruptions, but the present analysis only looks at the two types that represent the most severe forms of disruptions: *complete shutdowns* of all networks, and *partial shutdowns* that involve the shutdown of individual sites or subnetworks. The data range from 1995 to 2010, and within this time period offer details and circumstances of network shutdowns and disruptions in 101 countries (Howard, 2010: 222).

For the present analysis, I aggregate the events to the country-year level and match these to a panel dataset of all countries with a population larger than 500,000, in order to allow for a cross-national comparison between 1995 and 2010. Two disruption measures are extracted, hereby making use of the auxiliary information provided in the dataset. The first variable is a binary indicator that measures whether the government of a country implemented a full or partial shutdown of its digital networks in a given year (labelled as `full/partial disruptions dummy`). The second measure counts the number of full or partial disruptions by country and year (labelled as `full/partial disruptions count`).

[5]The dataset is provided by Howard (2011).

Figure 4.3 maps major government directed internet disruptions that occurred across the world between 1995 and 2010.[6] For the time period under consideration, network disruptions are a pervasive phenomenon across all parts of the world. The countries with the highest number of disruptions are China, Tunisia, Turkey, Iran, Vietnam, and India.[7] All of these countries have a history of repressing civil liberties, imprisoning political dissidents, and have struggled with armed internal opposition groups. Governments in the Middle East and North Africa made extensive use of internet disruptions during this time period, which pre-dates the start of the Arab Spring. The start of the region-wide uprisings have been traced back to the self-immolation of a Tunisian street vendor in December 2010, becoming a symbol of the civilian struggles across Arab countries. As the map shows, up until then, the Tunisian government had already implemented one of the most disruptive forms of network control globally.

A number of countries that are generally categorised as being democratic also display relatively frequent usage of network disruptions, such as Germany and Finland. Howard, Agarwal and Hussain (2011) explain that the majority of disruptions in democratic countries are justified by referring to legislation on child-pornography, and pornography in general. They describe the first disruption in Germany in 1995, where 'German prosecutors demanded that an ISP block 4 million worldwide subscribers from reading sex-related information' (Howard, Agarwal and Hussain, 2011: 219).

The absence of disruptive activities across large parts of Sub-Saharan Africa likely indicates the low levels of internet penetration across the region before 2010, prompting governments in this region to pay less attention to the opportunities and potential dangers of the internet. It is important to note, however, that low levels of internet penetration should not be taken as a sufficient indicator for the lack of government interest in surveilling and censoring its digitally connected population. The Ethiopian case study presented in Section 4.2 demonstrates that states where only a very small minority of the population is connected to the internet can still be highly motivated to heavily invest in surveillance tools and censoring methods. More generally speaking, the digitally connected subset of a country's population is likely to be comprised of the most affluent and educated individuals, an elite that is going to be critical for a government's longterm stability and survival (see Bueno de Mesquita et al., 2003). Consequently, governments will be particularly interested in controlling and surveilling this digital elite to foresee potential future instability.

Respect for physical integrity rights across countries and time

Comparing the respect for physical integrity rights across countries and time is a challenging exercise, as reporting practices vary substantially on both of these dimensions. The most accurate cross-national indicator currently available is the dynamic human rights scores by Fariss (2014), which are based on dynamic ordinal item response theory models (see also Schnakenberg and Fariss, 2014). The human rights scores estimate a latent measure of a country's level of respect

[6] The map was created using the cshapes and ggplot2 R packages (see Weidmann and Gleditsch, 2010; Weidmann, Kuse and Gleditsch, 2010; Wickham, 2009).

[7] Figure 1 in the Appendix provides a list of all countries highlighted in this map.

for physical integrity rights by combining information from a multitude of data sources capturing aspects of coercive state behaviour. These include the Cingranelli and Richards Human Rights Data (Cingranelli and Richards, 1999), the Political Terror Scales (Wood and Gibney, 2010), as well as specific datasets covering the prevalence of torture, genocide, and political executions (see Fariss, 2014: 302). Contrary to other human rights measures, the dynamic human rights scores allow the standards of reporting to vary over time, making the scores more comparable across the global sample, and over a period of time.

Common predictors of physical integrity rights protection

A number of standard variables that have been found to affect a government's willingness to enforce state-sanctioned violence are included (see, e.g. Poe and Tate, 1994; Poe, Tate and Keith, 1999; Davenport, 2007*a*). The presence of organised internal dissent is the most consistent and robust predictor for increases in state repression, and is measured using the UCDP/PRIO measure of armed internal conflict (Themnér and Wallensteen, 2013). An armed internal conflict is defined as an incompatibility that 'occurs between the government of a state and one or more internal opposition group(s) without intervention from other states' (UCDP/PRIO, 2014: 9), and that resulted in at least 25 battle-related deaths (this variable is named `civil conflict`). I include an additional binary variable that takes on a `1` where governments are involved in internal conflicts that resulted in more than 1,000 battle-related deaths in a given year (this variable is named `major conflict`), which is also based on Themnér and Wallensteen (2013).

To account for the regime type of a country, I make use of the Unified Democracy Scores by Pemstein, Meserve and Melton (2010), which provide a continuous variable with a comparable range to the human rights scores, where negative values indicate less democratic and positive values indicate more democratic institutions. Pemstein, Meserve and Melton (2010) combine the information provided by ten existing indicators for regime types and estimate a continuous measure using a Bayesian latent variable model, which is similar to the human rights scores by Fariss (2014). To account for size and wealth of a country, the population size, as well as the gross domestic product (GDP) per capita are included as control variables.[8]

Lastly, the internet has only recently become a platform for interactive, peer-to-peer communication. The earliest 'proclamation' of the so-called Web 2.0 goes back to 1999, where DiNucci (1999) contends that '[t]he Web will be understood not as screenfuls of texts and graphics but as a transport mechanism' (DiNucci, 1999: 221). The potential for collective organisation provided by this new form of digital communication is what governments are likely to fear the most. In an effort to account for the period of static internet usage versus the period of interactive usage, I include a binary variable that takes on a 0 for years up until 2000, and a 1 from 2000 onwards.

[8]Both measures are logarithmised to account for skewness, and are taken from Hunziker and Bormann (2013), who have attempted to correct previously misspecified calculations.

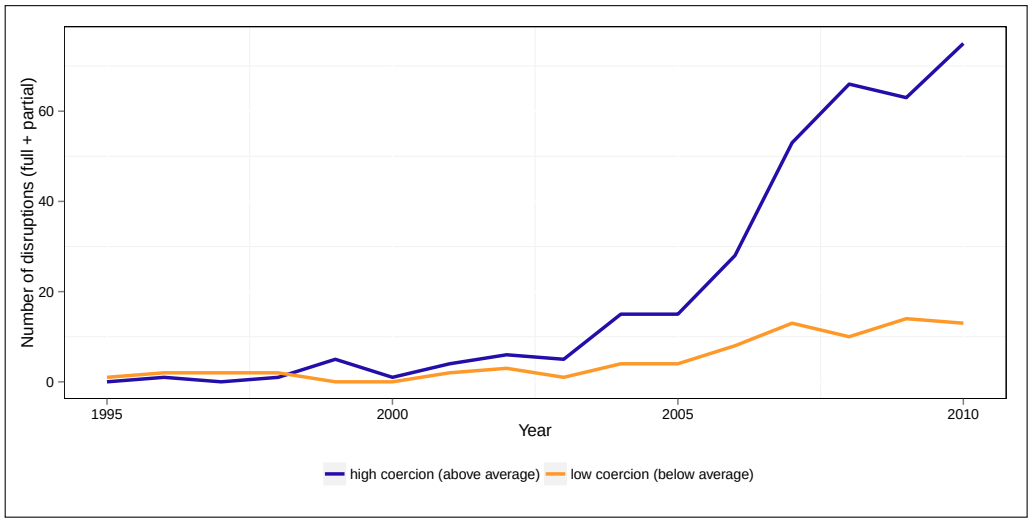


Figure 4.4. Government-directed internet disruptions, and human rights performance, 1995-2010.

Table 4.1. Summary statistics, a global analysis of network disruptions and violence state coercion

	N	Mean*	St. Dev.	Min	Max
coercion (latent mean)	2,721	0.458	1.322	−2.738	4.693
disruption (count)	2,738	0.152	0.911	0	31
disruption (yes/no)	2,738	0	0.272	0	1
civil conflict (yes/no)	2,721	0	0.336	0	1
major conflict (yes/no)	2,721	0	0.175	0	1
regime type (latent mean)	2,720	0.253	0.907	−2.023	2.263
log population	2,564	15.934	1.653	12.288	21.004
log GDP p.c.	2,564	8.486	1.373	4.764	11.980

* median displayed for categorical variables

Table 4.1 lists the summary statistics for the variables included in the analysis. The highest number of disruptions that were implemented in a country in one year between 1995 and 2010 was 31 incidences.

4.3.2 Results

Figure 4.4 presents a timeline of the number of network disruptions between 1995 and 2010. The two lines distinguish between disruptions implemented by governments that exercised a high level of violent coercion against their citizens, and disruptions by governments that exercised a low level of coercion. Since the human rights scores by Fariss (2014) are continuous, governments are categorised as highly coercive in years where they fall below the average level, and less coercive in years where they are above the average.

Descriptive Evidence

Figure 4.4 demonstrates that until the beginning of the 2000s, network disruptions were only rarely implemented in both rights-respecting and rights-abusing countries. From 2003 onwards, network disruptions are implemented at a higher frequency in general, but over the course of time until 2010, the number of disruptions in rights-abusing countries increases to become more than three times as high as the number enforced by rights-respecting countries. The latent measure of physical integrity rights violations adjusts for changes in reporting over time, and demonstrates that respect for these rights has in fact improved (Fariss, 2014). This means the substantial difference between rights-abusing countries making use of network disruptions and rights-respecting countries cannot be attributed to an artefact of the data.

The patterns presented in Figure 4.4 address a number of arguments put forward in this dissertation. First, network disruptions have exponentially increased since the 1990s, making them an issue of growing importance. Second, the exponential increase has principally occurred in countries with weak human rights records. Third, this increase commenced and has accelerated throughout the change from the static Web 1.0 to the dynamic, interactive Web 2.0, and the expansion of internet penetration across the world. Governments have evidently become increasingly aware of the collective potential of the internet, and fear the opportunities it offers its citizens to connect amongst themselves.

Multivariate Analysis

The descriptive evidence offers preliminary support for the notion that governments engaging in network disruptions are also more likely to abuse their citizens' basic human rights. There might, however, be a number of confounding factors that account for this observed difference. To address this concern, I estimate a number of multivariate panel models that include the variables most consistently found to explain human rights performance as control variables. I estimate both fixed-effects models, that account for unobserved heterogeneity between countries, and only look at variation within individual countries, and hierarchical random effects models. Table 4.2 presents the results, for the statistical model that accounts for civil conflict, whether there was a major conflict, the regime type, population size, wealth, and the binary post-2000 time variable. Models 1 and 3 specify network disruptions as a binary indicator (with Model 1 estimating a hierarchical model, and Model 3 a fixed effects model), whereas Models 2 and 4 specify them as the number of disruptions per year (again as a hierarchical and fixed effects specification, respectively). The results demonstrate that network disruptions, across different measurements and model specifications are significantly associated with lower levels of basic human rights respect. In years where states purposefully disrupted their internet, they were also significantly more likely to use state-sanctioned violence against their own population.

To ease the interpretation of the estimated effects, Figure 4.5 plots the point estimates and 95% confidence intervals of all explanatory variables in Models 2 and 4. Where the confidence interval includes 0 (highlighted by the dashed

Table 4.2. Network disruptions and state repression, random effects models

	Model 1	Model 2	Model 3	Model 4
Intercept	2.06*	1.95*		
	[1.02; 3.09]	[0.91; 2.99]		
Disruption	**−0.03***		**−0.03***	
(count)	[−0.05; −0.01]		[−0.05; −0.01]	
Disruption		**−0.12***		**−0.14***
(yes/no)		[−0.18; −0.07]		[−0.19; −0.08]
Civil conflict	−0.38*	−0.38*	−0.34*	−0.34*
	[−0.45; −0.31]	[−0.45; −0.31]	[−0.41; −0.27]	[−0.41; −0.27]
Major conflict	−0.16*	−0.16*	−0.15*	−0.15*
	[−0.26; −0.07]	[−0.26; −0.07]	[−0.24; −0.06]	[−0.24; −0.06]
Regime type	0.57*	0.57*	0.54*	0.54*
	[0.51; 0.63]	[0.51; 0.63]	[0.47; 0.61]	[0.47; 0.61]
Post 2000	0.07*	0.07*	0.01	0.01
	[0.04; 0.10]	[0.04; 0.10]	[−0.03; 0.05]	[−0.03; 0.05]
Log Pop.	−0.26*	−0.26*	0.43*	0.45*
	[−0.32; −0.21]	[−0.32; −0.20]	[0.22; 0.64]	[0.24; 0.66]
Log GDP p.c.	0.29*	0.30*	0.22*	0.23*
	[0.24; 0.35]	[0.24; 0.35]	[0.14; 0.30]	[0.15; 0.31]
Model	random effects	random effects	fixed effects	fixed effects
AIC	2385.07	2372.08		
BIC	2443.50	2430.51		
Log Likel.	-1182.54	-1176.04		
R^2			0.21	0.22
Adj. R^2			0.20	0.20
Deviance	2365.07	2352.08		
N	2548	2548	2548	2548
N (groups)	172	172		

*0 outside the 95% confidence interval

horizontal line), the effect is estimated to not be significant.[9] Figure 4.5 shows that even when controlling for these main confounding factors, between 1995 and 2010, governments that implemented network disruptions were also significantly more likely to exercise violent coercive force against their own population. Confirming results from previous analyses, states are more likely to abuse basic human rights when they are involved in a civil or a major conflict. Faced with credible challenges to their political authority, states will be highly motivated to increase coercive violence against their citizens in an attempt to regain their previous status quo. Furthermore, the regime type of a government significantly affects a government's inclination to employ violence domestically: the higher

[9]The coefficients can be compared with respect to their direction, but not with respect to their effect size, since the variables are not identically scaled.

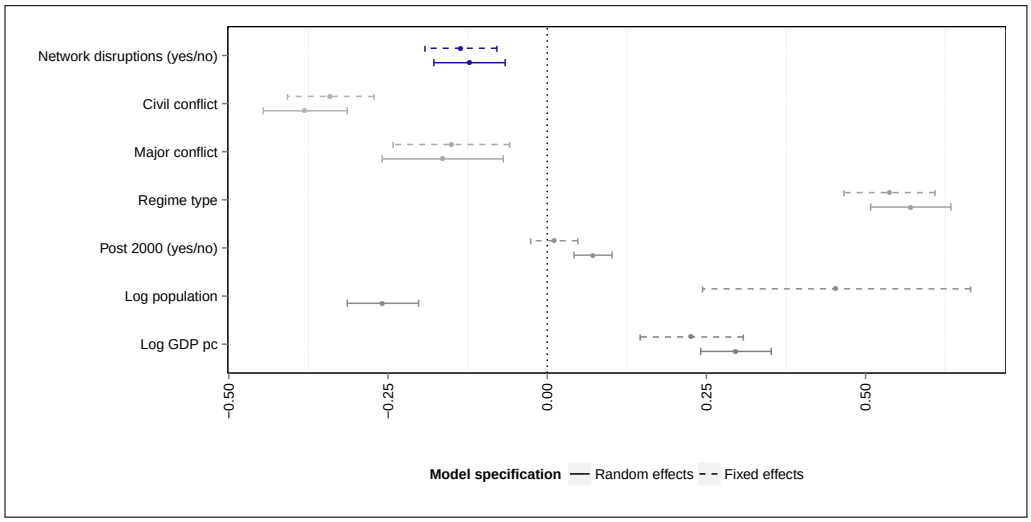


Figure 4.5. Network disruptions and repression, global analysis. Point estimates and 95% confidence intervals.

countries rank on the democracy scale, the more likely they are to respect and protect citizens' human rights.[10]

4.4 Summary

This chapter has presented case evidence and a global analysis for the argument that governments integrate varying forms of network control into their repressive strategies. The case evidence discussed in the first section of this chapter exemplified how governments choose between strategies of censorship and strategies of surveillance when faced with a perceived domestic threat, and how these strategies are then extended to the realm of coercion.

The second section of this chapter systematically investigated the prevalence and relevance of network disruptions with regard to human rights violations across the globe. The global analysis does not capture the trade-off states make when choosing between different repressive strategies, but it does offer substantial support for the relationship between internet disruptions and heightened state violence. State-implemented internet disruptions have heavily increased over time, and this increase can almost exclusively be attributed to governments that violate their citizens' basic human rights. The cross-national analysis from 1995 to 2010 showed that even when controlling for the most important factors that affect human rights respect, governments that disrupt their domestic internet are significantly more likely to abuse human rights.

To investigate the effect of varying forms of internet control on the nature of state-sanctioned violence, the remaining chapters of this dissertation move to the sub-national unit of analysis. The case under consideration is the Syrian regime's use of network control and strategic violence in the ongoing civil conflict. The next chapter presents a new database on state killings in the Syria that will form

[10]The population measure demonstrates the difference between the fixed effects and the hierarchical model: whereas highly-populated countries are more likely to abuse rights than less inhabited countries (random effects model), changes in population size within a country do not have this effect (fixed effects model).

the basis of the systematic subnational inquiries into how variations in network control affect strategies in violent coercion.

5

Integrated data on state killings in the Syrian Arab Republic

5.1 Introduction

To investigate the more fine-grained dynamics between network control and the nature of state violence, the theoretical and empirical investigations will now move on to the sub-national level and analyse how variations in network provision across the different Syrian governorates affects the regime's strategies of coercion. This chapter presents a new database on lethal state coercion between March 2011 and April 2014 in the Syrian Arab Republic. It discusses how the recent shift to disaggregated research on political violence makes significantly higher demands on the quality of the information needed to test empirical implications. Databases on violence tend to face two important challenges, which are the over-counting and the under-counting of incidences. I discuss how the first problem of over-counting violent events can be addressed through record-linkage, to ensure that all violent events are only counted once. The database on killings perpetrated by the Syrian regime that is presented here makes use of five different sources, of which at least one actively draws on media reports. The problem of duplicates in micro-level data on violence is particularly pertinent when relying on media-based sources, which form the basis for the majority of currently available datasets.

The new database on lethal state coercion in Syria therefore sets an important standard in that it not only identifies each documented lethal incidence by name, geographic location, date of death, and details on the circumstances of each death, but it also includes information on which and how many sources reported on every single event. The combination of geographic, temporal and situational information, paired with information on the density with which events were reported, makes the data suitable for more in-depth analyses of the patterns of violence committed by the Syrian regime, over the course of the ongoing civil conflict. The chapter concludes with a descriptive analysis of the data, and a discussion of the dark figure of violent events that failed to be reported at all.

Chapter 6 then presents one estimation solution to the problem of accounting for this dark figure, before Chapter 7 and 8 proceed incorporating the data and estimation strategy in their empirical analyses.

5.2 The problem of over-counting violence in event data

Research on political violence has benefitted greatly from a considerable increase in nuanced theorizing and empirical testing of how, when, where, and why people fight (Blattman and Miguel, 2010). Whereas important research on repression and conflict has focused on binary indicators, such as the absence or presence of peace (see Gleditsch, Nordkvelle and Strand, 2014), or ordinary measures, such as the Political Terror Scales (Poe, Tate and Keith, 1999), an expanding body of literature is now also interested in understanding the motivation for and effects of different types and intensities of violent patterns (Humphreys and Weinstein, 2006; Weinstein, 2007; Hultman, 2009; Stanton, 2009; Schutte and Weidmann, 2011; Sullivan, 2012).

Questions relating to the nature and dynamics of political violence are indisputably policy relevant, but their empirical testing will crucially hinge on the availability of high quality, micro-level information on violent incidences that offer a representative picture of the subject of interest. Fortunately, academic and human rights groups are doing their best to document violent events across the world (Raleigh et al., 2010; Salehyan et al., 2012; Sundberg and Melander, 2013), and the recent advances in automatic event coding are producing an ever-growing body of sources (King and Lowe, 2006; Schrodt and Idris, 2014). Nevertheless, collecting detailed information on violent incidences, in particular when perpetrated against civilians, is highly challenging, not least because it provide proof of serious crimes having been committed. Conflict actors and parties involved therefore have a general interest in hiding their atrocious behaviour and overstating the responsibility of their opponents (Slim, 2007).

It is important to note that the cause of these issues does not necessarily lie with data collection efforts, though different data projects may come to very different results even when attempting to cover similar grounds (see e.g., Eck, 2012; Chojnacki et al., 2012). The majority of causes for over-and under-counting violence are due to the nature of the underlying source material. Press releases, lists kept by human rights groups, and government or military reports all cover a certain snapshot of violent events. These snapshot views are aimed at fulfilling certain internal organizational goals. The majority of data on violent incidences remain dependent on journalistic sources, and ample research has demonstrated that in general, media agencies tend to pursue their own agendas. Usually, this does not involve providing a census of all violence occurring in contentious situations (see, e.g. Earl et al., 2004; Davenport, 2010). Furthermore, journalists working in highly unstable countries face extreme working conditions, and oftentimes have to deal with life-threatening danger to themselves (Arsenault, Himelfarb and Abbott, 2011). Changing security situations affect the accuracy and completeness of real-time reporting, and can lead to ex-post corrections of previously reported content.

Reflecting on the main problems of event data, which are generally based on news sources, Schrodt (2012) highlights the problem of duplicate stories as crucial in the case of both machine- and hand-coded data (for a comparison be-

tween these two types of data, see Hammond and Weidmann, 2014). Duplicated records can lead to serious over-counting of violent events. Schrodt (2012: 553) distinguishes between five different situations in which duplicate records might occur. First, smaller news agencies pick up the stories from large news agencies, such as Reuters, and these can get counted as separate events. Second, journalists might publish updates on the same violent incident as more information is revealed. Third, previous accounts of an incident are corrected with regard to the event size, date, and place, and these are mistaken for further events. Fourth, the same incident is referred to under differing headlines and published as parts of more general news briefings. Finally, different news agencies might report on the same event and focus on a different angle of the story. Schrodt (2012) concludes:

> 'when trying to measure trends in "ground-truth" behaviour against a baseline over a long period time, duplicates are a serious problem, both across sources and within sources. Cross-source duplication has probably changed considerably over the past 15 years due to local sources putting increasing amounts of material on the Web, and more generally with the globalization of the news economy [...]. In-source duplication can change due both to changes in the resources available to an organization [...] and editorial policies on updating, corrections, and the production of summaries.' (Schrodt, 2012: 553).

The database on lethal state coercion presented in the next section will therefore explicitly account for duplicate reports through record-linkage.

5.3 Data

5.3.1 Integrating multiple sources

To measure violent lethal coercion in the ongoing Syrian conflict, the data from five organizations that have been continuously working since the outset of the conflict, are cleaned and combined. It covers the time period from the beginning of the confrontation in March 2011 until end of April 2014. The data collection and integration process was conducted in collaboration with Megan Price and Patrick Ball as part of a report for the United Nations Office of the High Commissioner for Human Rights (Price, Gohdes and Ball, 2014). However, the composition of sources used in the report differs from the data presented in this dissertation in two important ways. First, the data presented in Price, Gohdes and Ball (2014) includes killings reported by the Syrian government. Since this dissertation focuses on victims killed *by* the government, the last-mentioned records are excluded in the present version of the data. Second, the data presented in Price, Gohdes and Ball (2014) does not include the records provided by the Syria Shuhada website, while these records *are* included in the present version of the data.

The five sources included in the analysis are the Syrian Center for Statistics and Research(SCSR)[1], the Syrian Network for Human Rights (SNHR)[2], the Syrian

[1] `http://csr-sy.org/`

[2] `http://www.syrianhr.org/`

Observatory for Human Rights (SOHR)[3], the Syria Shuhada (SS) Website[4], and the Violations Documentation Centre (VDC)[5]. While the records collected by Syria Shuhada consist of a combination of individually reported incidences, and reports from news sources, the other sources are all human rights groups working with local networks of informants on the ground, in order to obtain the most reliable information on the details and circumstances of those killed by the regime.

Defining lethal government coercion

The data used in this analysis include both combatant (such as belonging to the Free Syrian Army) and non-combatant victims killed by the Syrian government and pro-government forces. The data do not allow for an exact classification of victims into these two categories; instead they are classified as 'martyr' deaths by the recording groups, indicating that state military, paramilitary and other higher-ranking government officials are excluded. Since the object of inquiry in the present study pertains to the study of violence perpetrated by the government against whomever it deems threatening to its political stability, this can include ordinary civilians, and those who have mobilised an armed struggle against the government. From the position of the government, anyone who is not in active support of its regime is generally seen as a threat and treated as an anti-government combatant or collaborator, which is one of the main ways states justify the killing of non-armed citizens during episodes of civil conflict (Valentino, Huth and Balch-Lindsay, 2004; Slim, 2007).

5.3.2 Record-linkage

In order to assure the highest possible quality standards in combining documented evidence from different sources, records of fatalities are only included if they are identifiable by full name of the victim, date of death and governorate in which the death occurred.[6] The records are available at a daily level for each of the country's 14 governorates; further geographical disaggregation is not possible. All five data sources also provide auxiliary information on the circumstances of the death, but this information is not used in the record-linkage process.[7]

To create a complete and accurate list of documented killings these data need to be processed in two different ways: first, duplicates within individual lists have to be identified and removed. Fatality recording is conducted in the midst of chaos and fighting, making it highly probable that the same victim is recorded more than once by the same organization. This inflation of counts is likely to be non-random, as more visible attacks might lead an increased number of survivors to report the same victims. Second, victim identities need

[3] `http://syriahr.com/`

[4] `http://syrianshuhada.com/`

[5] `http://www.vdc-sy.org/`

[6] For further details see: 'Data Processing, Cleaning and Translation' in Price, Gohdes and Ball (2014: Appendix B).

[7] I use the details integrated from the different data sources to check the consistency of the linked records once the record-linkage has been completed.

to be linked across lists, in order to arrive at an overall number of documented victims.[8]

The identification of duplicates within a given source is termed de-duplication.[9] The identification of duplicates across different sources is termed record-linkage, or 'matching'. De-duplication and matching were completed in the same process by compiling one list that includes all records from all five sources, and searching for records that have the same information on the victim name, as well as place and date of death. Where information on the age, sex, and date and location of birth was available this was additionally used in the identifying process. For the period from March 2011 to April 2014, the present version of the data included 400,398 records from the five sources mentioned above into the record-linkage process.[10]

To facilitate the task of de-duplication and matching, the overall list of records was stratified by the governorate in which the individual was killed, as well as the year of death.[11] Each of these groups of records were then separately examined and searched for duplicate records. All records that were identified as referring to the same victim were then clustered. Each cluster thus represents one victim. For each cluster, the data sources that identified said victim were noted. The information on which data source identified which victim is a crucial part of the integration process that provides important information on the reporting process of the different organisations.[12] Table 5.1 presents a random sample of records from the integrated database to exemplify the structure.[13] Record 14567 was only recorded in one data source (in this case source 'A'), but record 78949 was reported in data sources C, D, and E. Each row in the integrated data base thus refers to one victim, but entails information on the origin of the information, and the number of sources that entailed this information. Note that the 'overlap' information provided in this format only tells us in how many different sources each record was found – it does not tell us how many times this record was found in one source, i.e. how many duplicates were found in one source. The reason for this is that the organisations providing the data all have different standards of pre-processing their data prior to sharing it. While duplicates within sources are problematic for further analyses, documenting their frequency holds no additional value, that could inform our understanding

[8] Records from the Syrian Observatory for Human Rights were not made available after April 2013. However, analysis of the matched data prior to May 2013 reveals that the contribution of records that are only identified by this one source is approximately five per cent, making the four source matching comparable that period where five sources were available.

[9] This description of the de-duplication and matching process is based on Price, Gohdes and Ball (2014: Appendix C).

[10] The data in Price, Gohdes and Ball (2014) includes records by the Syrian government, but excludes the data by Syria Shuhada, and reports an overall number of 318,910 records.

[11] In a second step of the process, cross-checking was conducted by comparing records from adjacent geographical locations, and records that noted a date of death at the beginning or end of a calendar year.

[12] If some of the auxiliary information on individual victims differs within a cluster, then the most frequently reported value was saved. For example, if an individual is identified in three databases by the same name, sex, and place of death, but two of the sources record the death to have occurred on the 22. April, and the third source reports the 23. April, then the data of death for this victim is recorded as the 22. April. In cases where two sources reported contradictory dates, the value was randomly chosen from the two available.

[13] Notice that the name is anonymised and the table does not report the auxiliary information on the circumstances of death.

Table 5.1. Snapshot of the database (anonymised)

ID	name*	A*	B*	C*	D*	E*	gov	date_of_death	sex
14567	[..]	1	0	0	0	0	Aleppo	2012-11-16	M
57860	[..]	0	0	1	1	1	Homs	2013-09-16	M
58673	[..]	0	0	1	0	1	Rural Damascus	2012-08-30	M
68900	[..]	0	1	0	0	0	Hama	2012-04-28	F
23456	[..]	0	1	0	1	1	Homs	2013-03-24	M
11239	[..]	0	0	1	0	1	Deir ez-Zor	2012-05-28	M
68900	[..]	0	0	0	0	1	Rural Damascus	2014-02-25	M
78949	[..]	0	0	1	1	1	Daraa	2013-03-18	n.a.

* anonymised

of where the data came from and how it was collected. The frequency of intra-source duplicates is therefore recorded but published in the final version of the integrated data.

The reliability the record-linkage procedure was checked by letting the coders all code the same sample of records, and then testing the inter-rater reliability (see Price, Gohdes and Ball, 2014: Appendix C1 for details). The coding conducted by the different human matchers was in agreement in over 97% of all records in the sample, indicating that the reviewers were highly consistent in their assessments of the records. Of the more than 400,000 records, 203,781 were identified as unique incidences of lethal violence.[14]

5.3.3 Descriptive comparisons

The frequency of reporting of individual records across different data sources can also be interpreted as the reporting *density*. The density of reporting can change over time, and it can vary across different locations. It can also vary across specific types of events or specific types of victims. The following graphs offer descriptive comparisons of the reporting density. Since one of the five sources, the Syrian Observatory for Human Rights, does not cover the entire period, it is excluded from these descriptive comparisons.

Figure 5.1 plots the de-duplicated number of records recorded by each sources over time, and compares them to the number of unique, matched records integrated from all sources. The fact that the integrated number of reported victims of lethal violence is higher than each individual source means that the different sources are contributing records that are not found in other sources – they are not mere replicates of each other. At a first glance, the overall trend shown by all lines looks fairly similar. Reported violence across Syria increases substantially in the summer of 2012, and spikes again in August 2013. The distance between the blue line (the integrated data) and the grey lines is, however, not constant. The number of unique records picked up by individual sources seems to increase substantially during the second year of the conflict.

To achieve a better understanding of the exact reporting density over time, Figure 5.2 shows the number of reported lethal violations for each month as bar

[14]The data in Price, Gohdes and Ball (2014) includes records by the Syrian government, but excludes the data by Syria Shuhada and reports 191,369 killings.

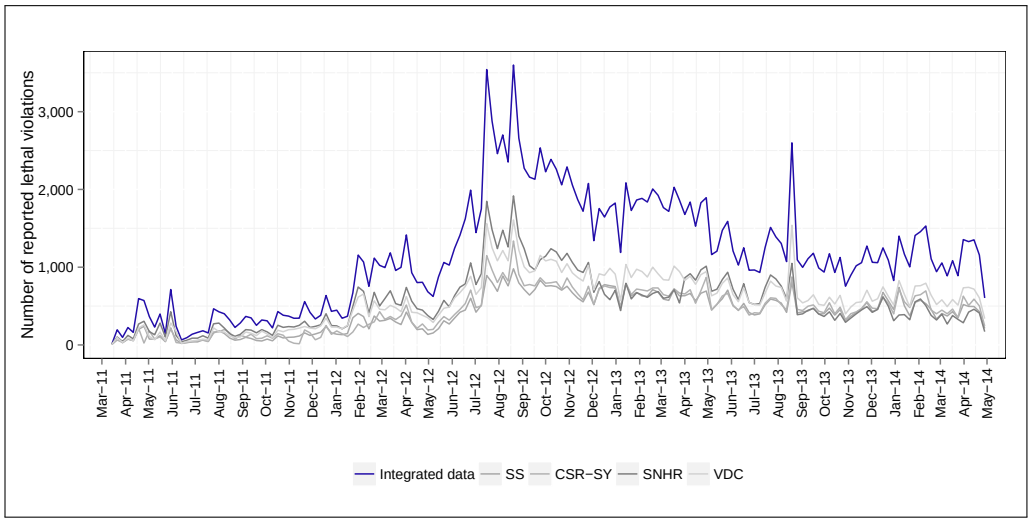


Figure 5.1. Individual sources, and integrated data, over time, Syria, March 2011 - April 2014.

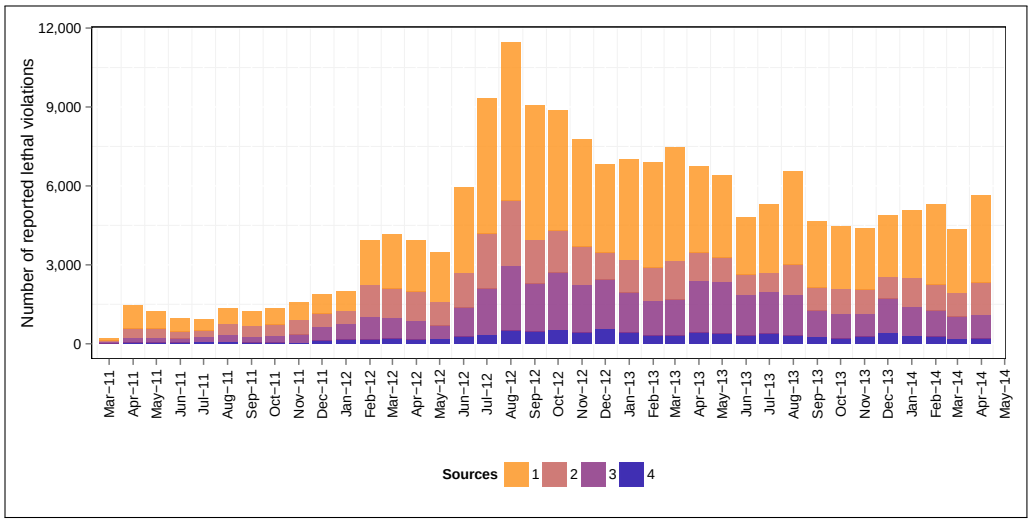


Figure 5.2. Density of reported monthly killings, Syria, March 2011 - April 2014.

graphs, but additionally shows how many of these violations were reported by one, two, three, or four of the data sources. The darker the bars are shaded in a given month, the higher the reporting density. Months where large parts of the bar is yellow indicate that many of the killed victims were only recorded by a single source.

Figure 5.3 breaks down the density of reporting for each Syrian governorate. Rural Damascus has by far the highest number of reported government killings in the period under consideration. Damascus and Deir ez-Zor have a similar number of killings reported in two, three, and four sources, but Damascus shows a larger number of killings that were only reported by one source.

A simple assumption might be that the reporting density of violent killings is related to the number of violent killings that occurred in the first place. This relationship is likely to be mediated by a number of important factors, such as the timing, type, and location where the events took place, but in general, it is

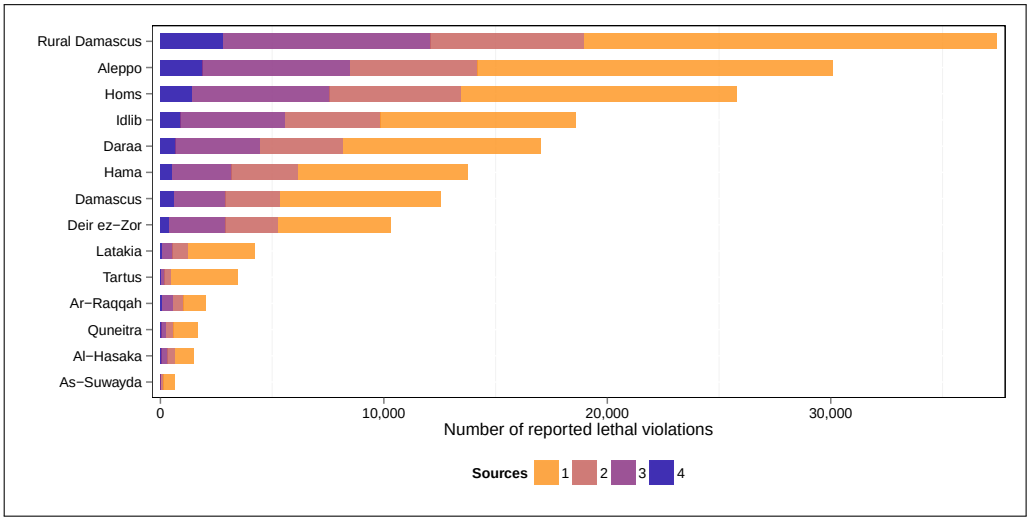


Figure 5.3. Density of reported killings, by governorate, Syria, March 2011 - April 2014.

plausible to assume that the fewer data sources report on a certain number of killings, the larger the number of killings that are missing. To give a related example, let us assume that a group of four government agents are given the task to provide a census of the number of protesters on a certain day in a capital city. The four agents walk all around the city individually, and each does her best to take a photograph of every protester. At the end of the day, the agents get together to compare their pictures. If all four agents took pictures of the same protesters, they can assume that the group of protesters is probably not much bigger than those captured by all of them. However, if they compare their pictures and realise that every agent has caught a significant number of protesters on camera that none of the others saw, they will realise that the crowd was much larger than any of them had anticipated. This will lead the conclusion that had they employed a fifth or sixth agent, the number of new pictures would probably have grown accordingly. In a similar way, the yellow sections of the bar charts in Figure 5.2 and Figure 5.3 indicate that a significant number of victims killed by the Syrian government have not been documented.

5.4 Summary

This chapter presented a new database on lethal violations perpetrated by regime forces in the ongoing Syrian conflict. To address a fundamental problem of event data on violence, which is the over-counting of individual records of those killed, the database integrates information gathered by five different human rights organisations working actively to document the atrocities being committed by the regime. Under the leadership of President Bashar Al-Assad, government forces have not spared the use of any gruesome actions to maintain political control.

The five sources were combined through record-linkage, a process by which duplicate records within each source, and across the different sources were identified and integrated. The linked database presents the (currently) most

accurate and complete list of documented state killings for the ongoing Syrian conflict. The fact that each recorded killing can be traced back to the original sources that documented it, makes it possible to investigate and model the variation of reporting density across different time periods and locations. The descriptive comparisons presented in this chapter show how the reporting density changes over the course of the conflict, and varies across different governorates. Although the temporal patterns of the individual sources look comparable at a first glace, the frequency of reporting changes significantly across time; this indicates that patterns of actual perpetrated state killings are likely to differ from the pattern of observed killings. While the integrated database tackles the problem of over-counting incidences in event data, it does not address the issue of unreported violence, or the under-counting of victims. This is the topic of the next chapter.

6

Accounting for the dark figure: unreported violence in event data

6.1 Introduction

Incomplete, unrepresentative data is a major problem for all research in the field of political violence. It is therefore important to discuss how this obstacle limits the inferences that can be drawn from empirical analyses based on the available data sources. I introduce a solution to this problem, showing how multiple-recapture models can be used to model the reporting process of multiple data sources on violence, and then predict the number of violent cases that went unreported. I demonstrate the reliability of the method by drawing biased convenience samples from a simulated population of violent incidences, and then predict how many cases were missed in the sampling process. The simulations show how the method picks up varying levels of 'missingness' across time and space, and works under conditions of both biased and unbiased reporting, providing the researcher with a tool for assessing the different datasets at her disposal. I then apply the method to reported state killings from the ongoing conflict in Syria, using the database presented in Chapter 5.

6.2 Information access and the challenge of identifying trends in event data

Research on the determinants and effects of political violence is generally interested in investigating variations in violent conflict behaviour. Some studies do this by looking at the absence or presence of violent episodes where a certain threshold of violence is fulfilled, such as 25 or 1000 battle deaths per year (for an overview, see Sambanis, 2004). More recent research has delved deeper into studying individual- or group-level dynamics of violence, thereby relying on information on violence that is disaggregated geographically, temporally, and by actors (Humphreys and Weinstein, 2006; Weinstein, 2007; Eck and Hultman, 2007). As established in the previous chapter, the majority of data sets build on

incomplete sources, such as for example media data, and consequently suffer from over-and under-reporting. While the question of over-reporting was addressed in the previous chapter, the question of under-reporting is the subject of this chapter.

Baum and Zhukov (forthcoming) analyse the media coverage of the recent revolution and conflict in Libya and find systematic evidence for both over- and under-reporting of specific kinds of protest and violent events, depending on the political context in which the media agencies are operating themselves. While media reports originating from non-democratic countries tended over-report the violence produced by the protesters, democratic countries disproportionately focused on state violence. The results presented by Baum and Zhukov (forthcoming) have two important implications for the use of media accounts as sources for event data. First, the Libyan revolution received high media attention around the world, not least because of the international intervention through NATO members, which had been mandated by the UN Security Council. When compared to instances of mass protest and state repression in other contexts, such as, for example in Sudan or Bangladesh, the coverage rates should be relatively high. Second, taking the numbers reported in media sources at face value means running the risk of counting individual events more than once *and* attributing them to the wrong perpetrator (see also Davenport, 2010).

Two recent studies by Weidmann tackle the problems of accuracy and bias in media-based conflict data for the case of Afghanistan (Weidmann, 2014*a*,*b*). Both studies compare the significant acts database of the US military (SIGACTS) to data collected by the Uppsala Conflict Data Programme on casualties in the Afghan war, which are spatially matched by events that occurred in 2008 and 2009. Analyzing the spatial accuracy of the media-based data, Weidmann (2014*b*) finds that press sources tend to report the locations of violent events more accurately if those events were larger in size, and if they occurred in more densely populated areas. With regard to bias – the incomplete reporting of violent events – Weidmann (2014*b*) presents simulations that show the damaging potential of incomplete data, which might fundamentally change the size and direction of estimated effects. For the case of Afghanistan, however, the levels of bias are found to be too feeble to affect results significantly, which can most likely be attributed to the strong international interest the conflict received.

Turning to a conflict that received significantly less media attention, Krüger (2014) matches and compares five data sources on violence during the Sierra Leone civil war, and finds that media-based databases report significantly fewer incidences across the entire conflict, and generally suffer from a 'capital bias', meaning that they under-report all events that occurred outside of the country's capital. She concludes that '[i]f five scholars were to conduct five case studies on the dynamics of violence in Sierra Leone during the last five years of the conflict, they would make different findings depending on which data source they chose' (Krüger, 2014: 45).

For studies that are concerned with understanding the substantial scale of violence in conflicts, a census of all violent events – that meet the operational definition of the research at hand – is required. Many theoretical questions do not necessarily require a census of all events, but need data that are representative of the census. For example, in investigating whether a certain violent actor has perpetrated fewer or more atrocities over the course of a conflict, the percentage

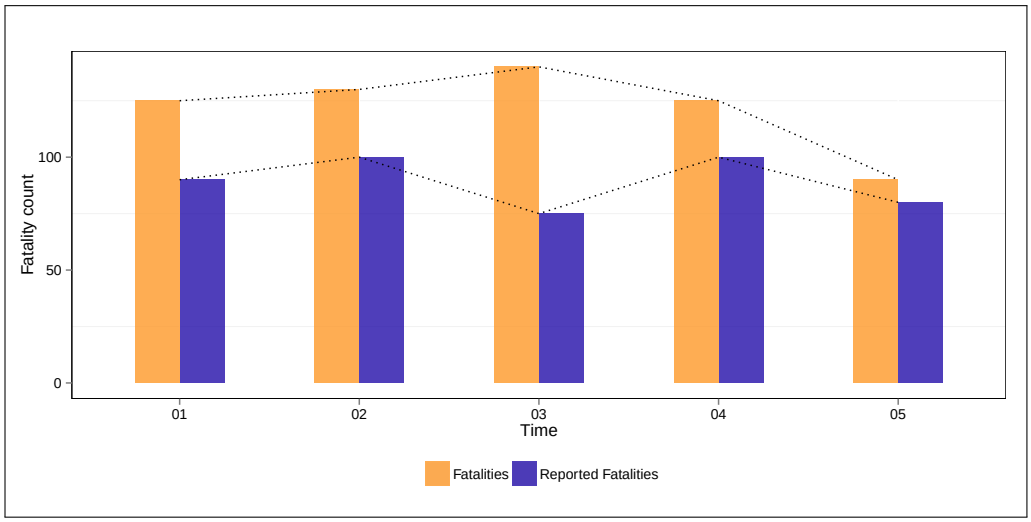


Figure 6.1. An example of violence and reported violence, over time.

of reported violence would have to remain relatively constant across time. Two problems arise from this: First, to obtain a representative sample of all violent events, researchers would require data that were generated with a probability-based sampling technique. In situations of extreme unrest and violence, this type of sampling is highly challenging to conduct. Retrospective mortality surveys that are conducted in the aftermath of conflict are another solution, but the associated costs and logistical challenges are substantial (Brück et al., 2010). Consequently, the second problem is that it is almost impossible to assess the representativeness of non-randomly sampled data without having access to the actual census. While meta-data and qualitative contextual knowledge on how the data were assembled helps the researcher assess potential causes of incompleteness, consistent representativeness across dimensions of interest to the researcher (such as time, space, and perpetrator) cannot be established.

Researchers are therefore frequently left with one or more data sources on violent events that tell incomplete, and sometimes even conflicting stories. The patterns produced by these data generally represent the reporting process, not the process by which violence itself was generated – in many cases leading researchers to draw inferences based on reported violence, not actual violence.

Figure 6.1 exemplifies how the problem of incomplete data can affect the inferential analysis of conflict dynamics. It shows a timeline of violence (the lighter bars) across five months, where the actual number of violent incidences steadily increases for three months, and then after that decreases. The reports of violence (the darker bars) mirror the increase in violence for the first two months, but then fail to cover the further increase in the third month. This type of change in reporting frequently happens when organisations lose some of their staff members, or when the situation on the ground turns particularly dangerous, or when the perpetrating groups change their warfare tactics from public displays to more clandestine operations. Following the month of reduced reporting, the number increases again and continues to follow the trend of actual events.

Without knowledge of the changed nature of reporting in month three, researchers using these data would be led to believe that violence decreased in the third month of the conflict, and that changes in covariates of interest might

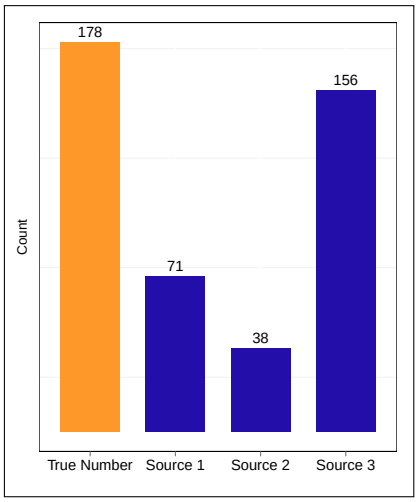


Figure 6.2. Violence, and three sources of reporting.

have been responsible for this. For example, if a rebel group changes its tactic from overt fighting to guerrilla tactics, this might lead to an actual increase in violence. If the groups reporting on the violence fail to immediately adapt their reporting procedure to this change, then the trend of reported violence is going to diverge from the actual trend. The inferences drawn from analysing these data would be incorrect. Modelling the process of reporting can offer a way to solve the problem.

6.3 Modelling the reporting of violence

One solution to getting at the number of victims that go unreported in conflicts is to study how those that *did* get included in NGO reports, government statistics, or press statements were recorded. In order to do this, two important pre-requisites need to be fulfilled: At least three different sources on the same conflict need to be available, and the incidences need to be recognizable across all available sources. In the majority of conflicts, a multitude of organizations are trying to keep track of events on the ground, although the motivation for doing so is likely to differ. Depending on each organization's partisanship, location, resources and mission, it will have a different propensity to record atrocities, and is likely to have access to a specific subset of the total population of victims.

If the number of reported violations can be modeled as a function of *who* was reported by *whom*, then the same model should be able to predict the number of violations that were not reported by *anyone*. The following example demonstrates this basic logic. Figure 6.2 shows a bar chart of reported and actual violence at a given place and time in conflict X. The entire number of killed individuals is 178, but not all of them made it into reports: 71 were registered by Source 1 (which might have been a morgue), 38 were recorded by Source 2 (which might have been a hospital trying to save these lives), and 156 were included in a report by Source 3 (which might have been an NGO). The highest number (the number 'closest' to the true value) is recorded by Source 3, but it is still failing to account for 22 victims. Those missing 22 victims might be included in the other two sources, but we do not know this (yet).

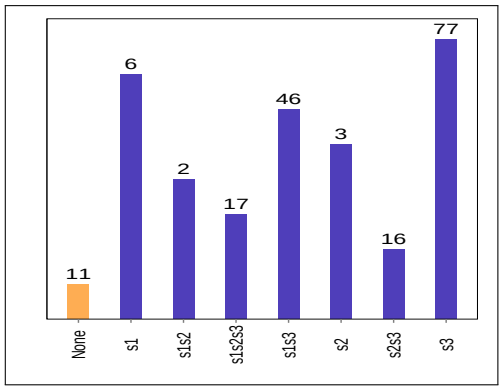


Figure 6.3. Reporting density across three sources.

Table 6.1. Example of individual and multiplicative capture histories

Y	x_1	x_2	x_3	x_{12}	x_{13}	x_{23}	x_{123}
17	1	1	1	1	1	1	1
6	1	0	0	0	0	0	0
2	1	1	0	1	0	0	0
46	1	0	1	0	1	0	0
77	0	0	1	0	0	0	0
16	0	1	1	0	0	1	0
3	0	1	0	0	0	0	0
??	0	0	0	0	0	0	0

Some victims are recorded by more than one source. The fact that different sources tend to have overlapping information is the main reason why it is not possible to just sum up the numbers of the individual sources: this would lead to certain victims being counted more than once. In order to model the reporting process, it is crucial to distinguish between the different reporting groups. Figure 6.3 graphically depicts the overlaps: When all three sources are matched against each other (and compared to the actual number of killed), 11 victims are found to be missing from all three sources. Six are only reported by Source 1 (labeled $s1$ in the figure), 2 are reported by Source 1 and 2 (labeled $s1s2$), 17 were reported by all three sources (labeled $s1s2s3$), etc. In a real world case, the number of unreported victims is not known to the researcher, but the reporting overlap, or density of reporting can frequently be extracted from the reported data. Estimating a census from an incomplete contingency table (which is what the reporting overlap provides) has been long discussed in the field of biostatistics, and log-linear capture-recapture models offer a powerful and flexible formalization for predicting the 'missing cell' of unreported violence (Cormack, 1989; Fienberg, 1972). Different formalizations of the method have been used to project populations of violent incidences in Kosovo, Peru, Timor-Leste, Colombia, El Salvador, and Guatemala (Lum, Price and Banks, 2013; Lum et al., 2010; Ball et al., 2003, 2002; Ball, Kobrak and Spirer, 1999; Hoover Green, 2011; Manrique-Vallier, Price and Gohdes, 2013).

In order to accurately model the reporting process of conflict violence, potential dependencies between the sources need to be taken into account. Three different types of dependencies can be distinguished:

Positive dependencies

Some sources are more likely to cover the same sample of victims than others. For example: media sources might be more likely to report on large, visible events where the number of people killed was high. Or two human rights groups contact the same people in rural regions in order to update their records of who is being killed. Some groups might even work together to reduce their personal risks where heavy fighting is occurring. Further dependencies might arise from victim characteristics: middle-aged men and women might be recorded more

frequently than the deaths of elderly or young victims. Reasons for positive dependencies are myriad, and they all lead to the same outcome of having a larger overlap in reported violence between two (or more) sources.

Negative dependencies

Political affiliations are likely to affect which victims end up being recorded by which source. Groups that receive support by the government are more likely to concentrate on victims killed by the challenger, and witnesses are more likely to trust these groups if they are reporting a rebel-perpetrated killing. Groups on the other side of a frontline might correspondingly have more access to information on regime-perpetrated victims. Other types of negative dependencies might be related to other victim characteristics, such as the gender, religion, age, or ethnicity of the victims. Groups with negative dependencies are likely to have less overlap in their data.

No dependencies

Two or more sources are independent in situations where all victims have the same probability of being included in each source. For example, if two organizations administered random sampling techniques, one could assume their reporting processes to be independent of each other.

To model these dependencies, log-linear capture-recapture estimation effectively formalizes these interactions as additive multiplicative terms (Bishop, Fienberg and Holland, 1975).[1] It assumes that Y, (the overlaps, or reporting histories) can be modeled as

$$\log(Y) = \alpha + \beta X + \epsilon \tag{6.1}$$

where X is a matrix of binary variables indicating the reporting history of each overlap (x_i, x_j, x_k), and multiplicative dependencies between the sources (x_{ij}, x_{jk}, x_{ijk}, etc.). Table 6.1 shows an example for such capture histories for three sources, but the model is extendable to more sources.

Following the three-source example, there are eight different ways to model the dependencies of the reporting processes. The first model assumes no dependencies between the sources:

$$\log(Y) = \alpha + \beta * x_1 + \beta * x_2 + \beta * x_3 \tag{6.2}$$

Second, there are three models that assume two lists are dependent:

$$\log(Y) = \alpha + \beta * x_1 + \beta * x_2 + \beta * x_3 + \beta * x_1 x_2 \tag{6.3}$$

$$\log(Y) = \alpha + \beta * x_1 + \beta * x_2 + \beta * x_3 + \beta * x_2 x_3 \tag{6.4}$$

$$\log(Y) = \alpha + \beta * x_1 + \beta * x_2 + \beta * x_3 + \beta * x_1 x_3 \tag{6.5}$$

[1] Log-linear multiple recapture analyses can easily be implemented in the R package `Rcapture`, which was written by Baillargeon and Rivest (2007).

Third, there are three further models that account for dependence between two pairs of lists:

$$\log(Y) = \alpha + \beta * x_1 + \beta * x_2 + \beta * x_3 + \beta * x_1x_2 + \beta * x_2x_3 \tag{6.6}$$

$$\log(Y) = \alpha + \beta * x_1 + \beta * x_2 + \beta * x_3 + \beta * x_2x_3 + \beta * x_1x_3 \tag{6.7}$$

$$\log(Y) = \alpha + \beta * x_1 + \beta * x_2 + \beta * x_3 + \beta * x_1x_2 + \beta * x_1x_3 \tag{6.8}$$

Lastly, there is one model (the saturated model), accounting for dependence between all lists:

$$\log(Y) = \alpha + \beta * x_1 + \beta * x_2 + \beta * x_3 + \beta * x_1x_2 + \beta * x_1x_3 + \beta * x_2x_3 + \beta * x_1x_2x_3. \tag{6.9}$$

Since Y represents a contingency table, it is usually modeled as a Poisson distribution. To select the simplest model that best fits the reporting process of the data, all models are estimated, and then assessed with respect to their goodness of fit, generally determined through the BIC.

For the example presented in Figure 6.2, the model that best fits the reporting process is the independent model, with:

$$\log(Y) = 2.276 - 0.398 * x_1 - 1.295 * x_2 + 2.018 * x_3 \tag{6.10}$$

Since Source 3 records the largest number, it is not surprising that the coefficient for x_3 is large and positive.

The fitted reporting model is used to predict the number of unreported incidences by setting all $X = 0$:

$$\log(Y_{000}) = \alpha, \tag{6.11}$$

so that $exp(\alpha)$ predicts the number of incidences when all X are zero. For the just-mentioned example the exponential of the intercept is approximately 10. Based on the best fitting model of reporting, the predicted number of incidences that where missed is 10 (with a 95 confidence interval [3:18]). As shown in Figure 6.3, the 'true' number missing is 11. The predicted population of violence is therefore the number of observed + predicted unobserved, which in this case would be 177 (with the true number being 178).

6.4 Simulations

To demonstrate more generally the potential of this basic method of prediction, I simulate a population, or 'context' of violent incidences that have two defining characteristics, which are the time and place where they occurred. Figure 6.4 visualizes the simulated population, which varies across eight time points, and two locations – defined as urban and rural. In total, the population includes 5000 incidences that are distributed along these two variables of interest. Next, I construct three types of hypothetical sources that collect incomplete information from this population. The probability of each incident being included differs between the sources, and is determined by the parameters of simple logistic equation for each source:

- $p_{S1_tp} = \text{logit}^{-1}(.1 + .5x_t - 3x_p)$

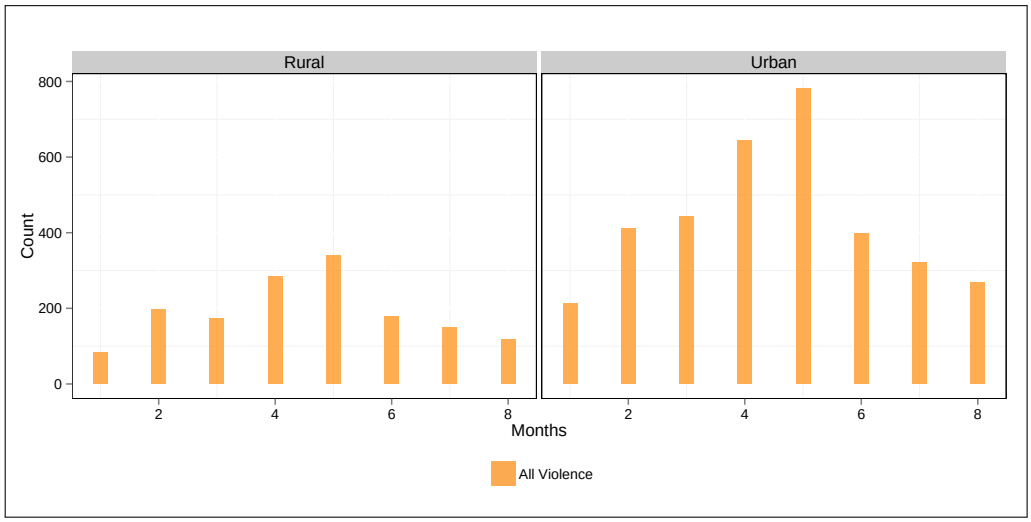


Figure 6.4. Simulated levels of violence in urban and rural regions.

- $p_{S2_tp} = \text{logit}^{-1}(.5 - .3x_t - .5x_p)$
- $p_{S3_tp} = \text{logit}^{-1}(-.2 - .001x_t + 2x_p)$,

where p_{Si_tp} is the probability of a violent incident being included in Source i at time t, and place p. The probabilities are dependent on a constant, the time the violence occurred (x_t), and whether it occurred in an urban or a rural location (x_p). I simulate 1000 rounds for each source, drawing the samples from a binomial distribution using the probabilities defined by the logistic process. Figure 6.5 presents one round of simulated sources by time and place of violence. When compared to the Figure 6.4, it is clear that all sources are missing at least some information at some point, but that the degree of completeness differs. Source 3 samples significantly more incidences that occurred in the rural than the urban location. Analyses based on Source 3 would therefore considerably over-estimate the proportion of violent incidences in the rural location. Conversely, Source 1 captures a larger sample of urban incidences. Source 2 changes its propensity to record incidences over time, providing a very different temporal trend to the one in the actual population. Since the degree of bias (and completeness) varies significantly across time and location in all sources, the process of reporting is likely to vary by time and place as well. For each of the 1000 rounds of source sampling, the reporting process is modeled, and the level of underreporting predicted for each time and place individually, producing 1000 estimates of underreporting for every time and place combination, or stratum. Figure 6.6 visualizes the results of the simulations for every stratum. The reported samples of violence by Source 1, 2, and 3 are depicted in shades of blue. The estimated number of actual violations is depicted in yellow, and the true number (in the simulated population) is marked at the red line.

The simulation demonstrates that the predicted number of actual killings converges with the true number defined in the population. Importantly, it is designed to test how the method weathers under different sampling conditions – all of which reflect probable real-world scenarios. Terrorist attacks in capital cities are likely to be closely scrutinised and documented by organizations

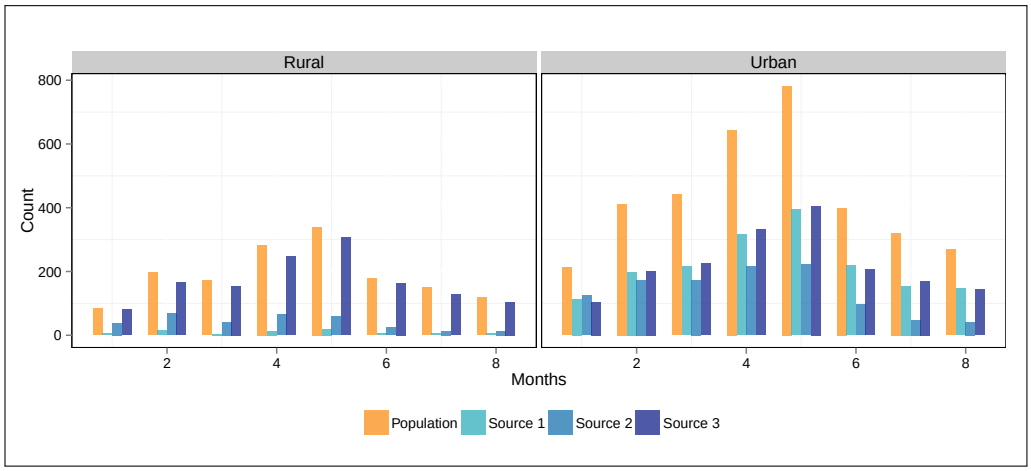


Figure 6.5. Example of simulated reporting of violence in urban and rural locations.

present, whereas the death tolls produced by clandestine operations in the countryside are likely to only partly show up in official statistics. For example, the top left graph in Figure 6.6 shows the samples and estimates for all killings that occurred at $t = 1$ and in urban locations $p = urban$. All three sources underreport violence by at least 50 incidences; nevertheless the predicted number converges with the true value. In other cases, such as at $t = 7, p = urban$, one source reports almost all incidences. In such cases, the predicted number of cases equally converges with the true number. Since the method works for different degrees of incompleteness and bias of the source data, it can be used to correct for missing cases *and* to check the consistency of the data over different dimensions of interest.

In the following section, I demonstrate how the method can be used to predict recent trends in conflict violence in the ongoing Syrian conflict. The analysis shows how the scale and trend of violence differ profoundly between predicted actual levels of violence and what was observed.

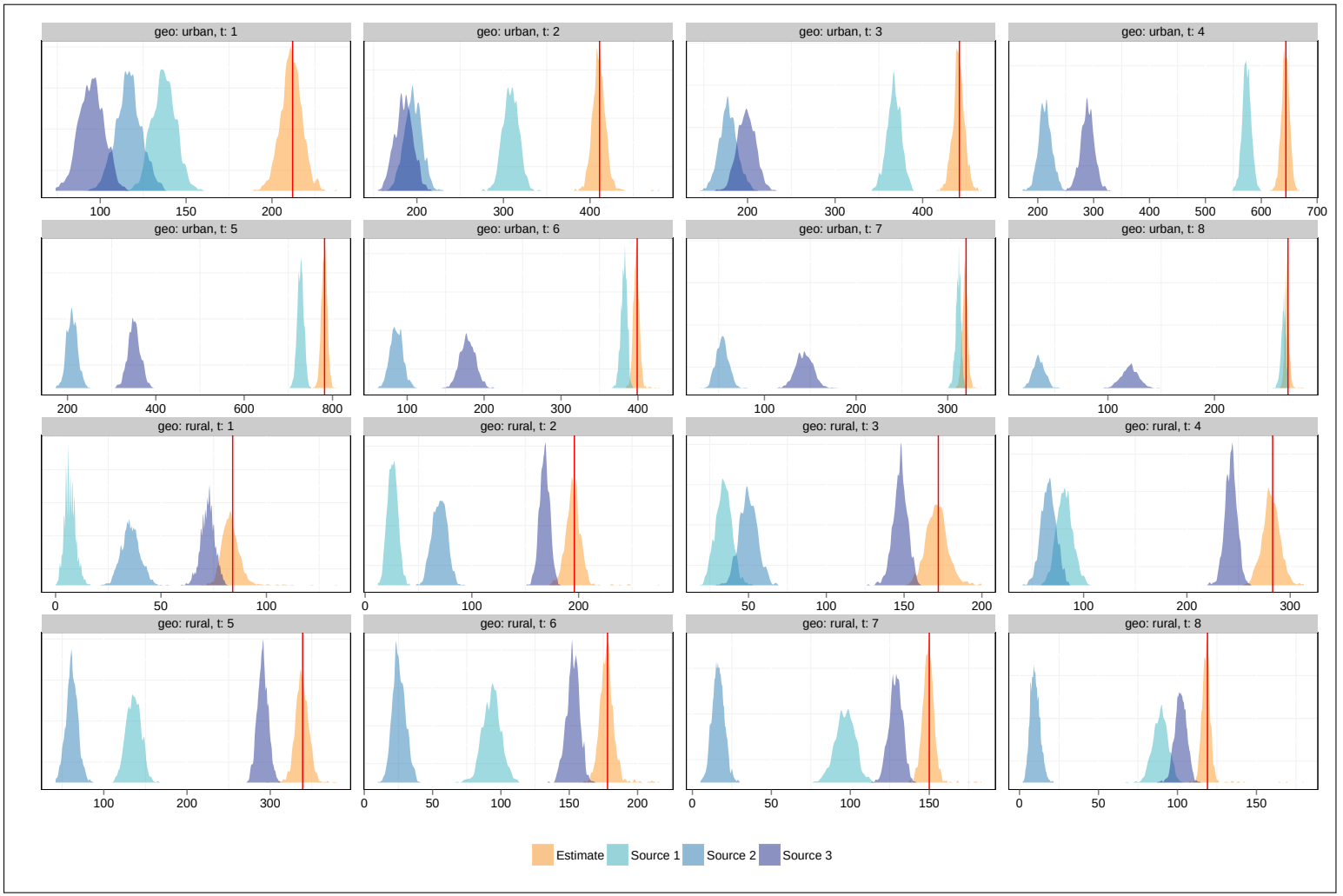


Figure 6.6. Simulations: reported and estimated violence.

6.5 An example: control, contestation and violence in Syria

The following application is intended to demonstrate how the use of incomplete, unrepresentative event data might affect an analysis of violent dynamics. The example chosen here is not directly related to the theoretical argument put forward in this dissertation; it does, however, make partial use of the new data on state killings in Syria.

The Syrian civil war has just entered its fourth year and has forced millions of Syrian civilians to flee their country to seek safety from the campaign of violence that has been erupting between the regime led by Bashar Al-Assad and a number of different opposition groups (see Lynch, 2013).

A central theoretical claim that has been put forward in recent research on civilian victimization relates the severity of conflict to the level of territorial control in a given location (Kalyvas, 2006). Kalyvas' theory of selective and indiscriminate violence contends that the degree of territorial contestation between two conflict parties principally determines to what extent perpetrators will only kill selected civilian targets, and to what extent their strategy of warfare will be indiscriminately atrocious against innocent bystanders. In essence, he contends that higher levels of control lead to increased information about who is supporting whom, which makes it easier for the warring party in control to choose their targets selectively. Conversely, where control is contested, the conflict parties have a harder time identifying their enemies, and civilians have less incentives to denounce others, as the benefits are unclear where control is unclear (see Kalyvas, 2006; Kalyvas and Kocher, 2009; Kocher, Pepinsky and Kalyvas, 2011; Bhavnani, Miodownik and Choi, 2011). Territorially contested regions are therefore expected to produce substantially higher numbers of indiscriminate killings than regions controlled by one conflict party. Focussing only on the level of violence perpetrated by the government, the following empirical expectation can be formulated to exemplify an analysis of violent dynamics in civil conflict:

- *Empirical Expectation:* Government forces will perpetrate higher levels of violence where they do not possess full control of a territory.

To what extent is the level of civilian victimization in the current Syrian civil war a function of territorial contestation? To study the impact of contestation on the level of violence, I analyze two central locations of violence within Syria that have witnessed different developments of territorial control over the course of the conflict. At the beginning of 2012, both Damascus and Aleppo were under full territorial control of the Syrian Regime. In both locations, rebel forces challenged the government's control in mid-July 2012. In Damascus, the government successfully regained control of the majority of the city within less than three weeks. In Aleppo, the contestation between different opposition groups, including the Free Syrian Army and the Islamic Front, has continued since its start in July 2012, and the city remains divided between the different conflict parties (Holliday, 2013).

Based on the theory of territorial control and civilian victimization, the empirical expectation would be that Aleppo and Damascus should *both* have witnessed an increase in indiscriminate killings in mid-July 2012. Given the quick regain of territory in Damascus by the government, the expectation would then diverge between the two locations: we would expect to see a persistently

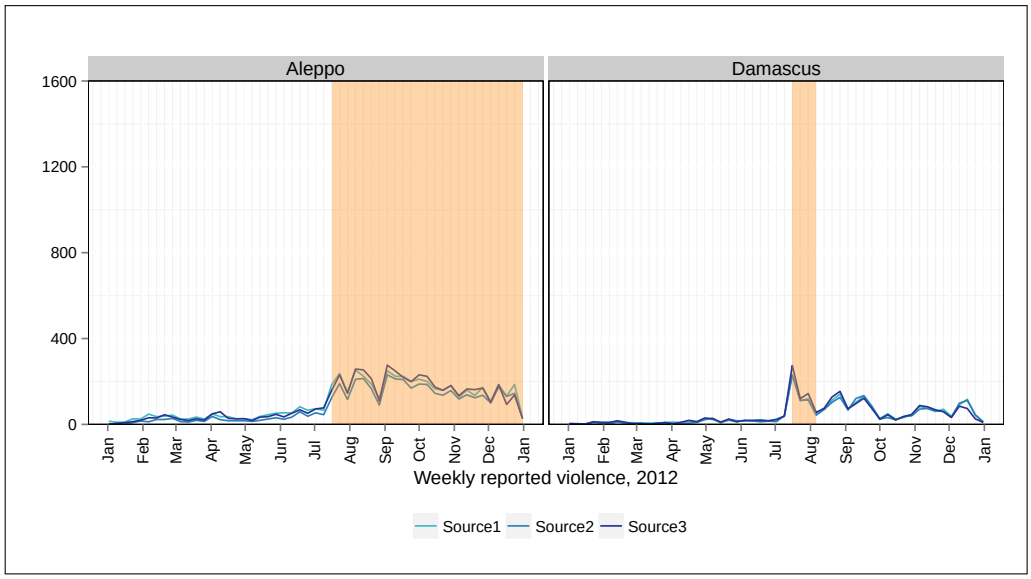


Figure 6.7. Weekly reported violence in Aleppo & Damascus, 2012.

high level of violence in Aleppo, whereas the number of fatalities should have decreased to fewer and targeted individuals in Damascus.

Figure 6.7 shows the number of weekly de-duplicated reported state killings that occurred in Damascus and Aleppo in 2012, and were collected by three well-known sources, that are included in the integrated database presented in Chapter 5.[2] The sources show very similar conflict dynamics up until mid-July. For both locations, the reports of violence show a marked increase in mid-July (marked by the beginning of the yellow-shaded segment), when the uprisings began in both localities. Reported violence, however, remains at a higher level in Aleppo than reports of violence in Damascus. These patters offer preliminary, descriptive support for the theoretical expectations – Aleppo's contested territory remains more violent than government-controlled Damascus.

But what about the data generating process of reporting? Might the reporting patterns differ between Damascus and Aleppo, and do they change over the course of the year 2012? Using multiple recapture modelling, I predict the number of unreported incidences of violence by week, in order to investigate whether the reporting patterns differ from estimated levels of conflict violence. Figure 6.8 shows the same timelines as Figure 6.7, but includes the estimates of projected weekly violence across time in Aleppo and Damascus.

Two important findings emerge. First, the change in territorial control affects the reporting process in *both* locations: after the start of the uprisings in mid-July both Damascus and Aleppo witness far higher levels of violence than were reported. Second, with regard to the empirical testing of the theory, the level of violence in government-controlled Damascus is predicted to have remained at a substantially higher level than reported. Evidently violence did not decrease to

[2]For demonstration purposes I only present this analysis with three sources in this chapter, but the results using four sources were essentially the same. The sources included in this analysis are the Syrian Network for Human Rights (SNHR: `http://www.syrianhr.org/`), the Syrian Center for Statistics and Research (SCSR: `http://www.csr-sy.org/`), and the Violations Documentation Centre (VDC: `http://www.vdc-sy.org/`).

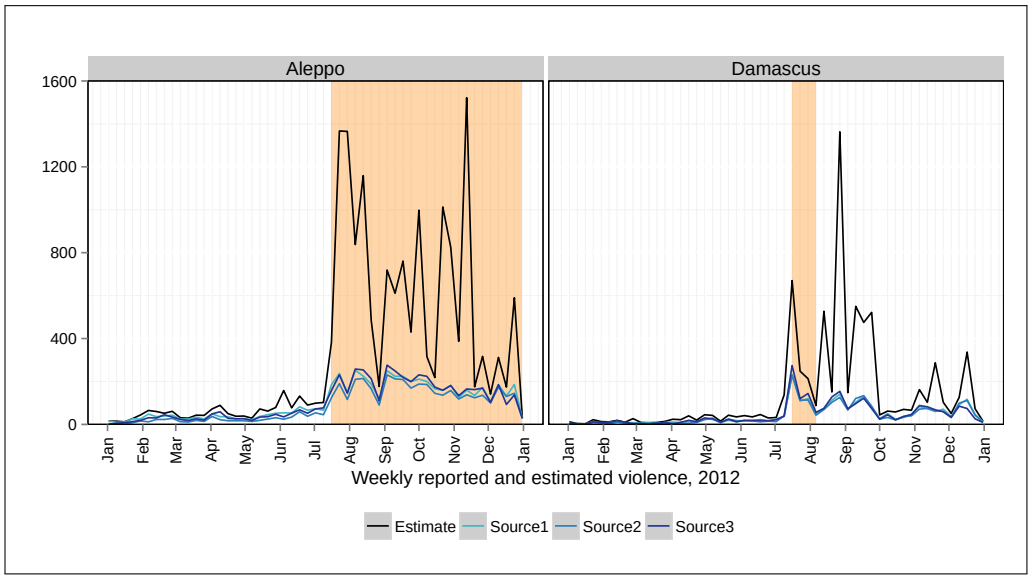


Figure 6.8. Reported and predicted violence in Aleppo & Damascus, 2012.

pre-uprising levels in Damascus, even though the government regained control of the city.

To analyze the effect of the uprising on the change in violence, I estimate a time series count model for Aleppo and Damascus, and replicate the model for all three sources, as well as the estimates. To account for the dependencies across time, I use the Poisson Exponentially Weighted Moving Average model by Brandt et al. (2000), where the level of violence at time t is modeled as a moving average of previous violence. For both Aleppo and Damascus I include a 'Battle' dummy variable that takes on a 1 in the weeks where the territory was contested, and a 0 otherwise. Table 6.2 reports the predicted percentage changes in violence during periods of contestation, as compared to previous government control. For Damascus, I include a 'post-battle' dummy variable that takes on a 1 for the period *after* the uprising, and a 0 otherwise. For Aleppo,

Table 6.2. Predicted percentage changes of violence during battle

Data	Battle (Aleppo)	Battle (Damascus)	Post-battle (Damascus)
Estimate	467.9 %	640.5 %	242.7 %
Source 1	182.4 %	547.8 %	191.7 %
Source 2	188.9 %	572.8 %	181 %
Source 3	125 %	640.5 %	212.7 %

No Post-battle effect changes for Aleppo reported, as it remained contested throughout 2012.

the model with the estimated level of violence predicts that during the period of territorial contestation, Aleppo fell victim to an increase in violence of almost 470%. The three individual sources all clearly predict a far lower increase of merely 120-190%. The reported data therefore significantly underestimates the change in the level of violence that occurred when the level of territorial control changed.

In Damascus, the change in violence throughout the period of territorial contestation is predicted to be even higher than in Aleppo, with an increase of over 640%. However, in the case of Damascus, Source 3 provides the same predicted change as the model with the estimated level of violence shows. Evidently, in some instances, the dynamics of violence are reflected correctly in reported data. When looking at the predicted change in violence in the aftermath of territorial contestation, we again see that the reported data underestimates the level of violence.

The results offer important implications: They provide substantial empirical support for the fact that changes in territorial contestation are likely to lead to higher levels of civilian victimization. Importantly, the findings reveal that reported data runs the risk of not being able to 'keep up' with recording incidences in times of intense conflict, leading to empirical implications that underestimate the actual effects of fighting. Furthermore, all data sources included in the analysis publicly support the Syrian opposition. It is plausible to assume that they face increased surveillance and additional obstacles in recording casualties in areas controlled by the government, which might explain why so many fatalities went unreported in the aftermath of the rebel uprisings in Damascus in the summer of 2012. Modelling the reporting process and using it to predict more reliable numbers reveals these challenges and shortcomings, which would otherwise have gone unnoticed.

6.6 Summary

Understanding patterns of atrocious state behaviour is a crucial part of disaggregated research on political violence, which has gained increased interest in recent years. Too often, however, empirical tests of theoretical explanations suffer from incomplete and biased measures of violence.

I discuss a well-known solution to estimating hard-to-reach populations when the only data available are convenience samples. Multiple-recapture methods were first developed to estimate wildlife populations and have since found their way into research on demography, epidemiology, and casualty estimation (for a review, see International Working Group for Disease Monitoring and Forecasting, 1995*a*,*b*; Manrique-Vallier, Price and Gohdes, 2013). Where at least three data sources are available, I show that a model of the reporting process of violence can be used to predict unreported incidences. To illustrate the method, I draw multiple biased samples from a population with two covariates, model the reporting processes and use this to predict the population size. The simulations using artifactual data demonstrate that the population size is, on average, predicted correctly both in cases where the samples are close to the actual population, and in cases where the samples are extremely biased. I then apply the method to data on violence collected in the Syrian civil war in 2012, and show how the predicted patterns of violence differ from the observed patterns.

To date, researchers have not made use of the results obtained from multiple recapture estimation to correct potentially biased data on violence used in conflict research. The implication is that the use of incomplete data for empirical analysis, if not collected through a sampling method that guarantees representativeness, runs the risk of producing statistical estimates that describe the pattern

of reported violence, not of the actual dynamics. In the next chapter, I make use of this estimation strategy to compare variations in the reporting of state killings prior to, and during, state-implemented internet outages.

7

The military strategic value of national network disruptions

7.1 Introduction

Governments fighting to maintain political control have an incentive to implement internet blackouts in conjunction with larger military offensives aimed at restoring control. Regime forces are likely to utilize these shutdowns as a tactical advantage when facing intense confrontation from violent opposition groups. The reduced opportunities for short-term military coordination of attacks is expected to improve government-aligned fighters' chances of regaining control previously challenged by anti-government fighters. If the shutdown of all communication networks is implemented when repressively responding to increased resistance, regime forces are likely to be involved in increased fighting directly prior to, and during the period of the outage. One observable consequence is therefore constituted by an increase in violence perpetrated by regime supporters immediately prior to, and during such outages.

I empirically test this proposition using new data on reported daily incidences of fatal regime violence in Syria, presented in chapter 5. I find that government-induced network blackouts are accompanied by significantly higher levels of violence, in particular in the governorates where government and opposition forces are directly confronting each other.

An alternative explanation might be that governments do not anticipate operational advantages by cutting connections, but instead implement blackouts to commit atrocities that are hidden from international scrutiny. To test for this cover-up hypothesis, I use log-linear capture-recapture models to estimate the degree of underreporting of conflict fatalities during blackouts, and compare it to the already existing levels of underreporting on days prior to network outages. The evidence suggests that unreported violence does not systematically increase during network outages, which is most likely attributable to the very short disruption intervals. Instead, the increase in documented violence indicates

that network outages are likely to form a part of the Syrian regime's coercive response strategy.

In the following two sections, recent research on censorship and blocking of the internet as well as the relationship between network accessibility and the potential for conflict is reviewed. I then discuss the theoretical motivations and potential costs for governments who disrupt network services while being challenged by armed opposition groups. Following a discussion of possible alternative motivations, I formulate empirical expectations that can be tested to uncover the predominant motivation behind implementing outages. The empirical section introduces the data on regime violence and network outages in Syria, and proceeds with the results of the analysis of documented violence, and the variation in documentation patterns during outages. The chapter concludes with a discussion of the results and potential avenues for future research.

7.2 Outages and operational advantages

The theoretical framework in Chapter 3 clearly shows that governments intent on using modern communication technologies to their advantage in repressing opposition groups have developed a resourceful set of tools to do so without shutting down all virtual and mobile access for their population. Why then, do so many countries experience purposefully implemented shutdowns? In theory, there are two types of access denial: the long-term prohibition or restriction of internet access, and shorter, infrequent intermissions of accessibility. As discussed in Chapter 2, the literature on censorship of the internet indicates that content censoring usually happens over longer periods of time, and is principally aimed at repressing collective organization and mobilization of disgruntled citizens (Howard, Agarwal and Hussain, 2011; King, Pan and Roberts, 2013). It is therefore plausible to assume that long-term outages implemented by the government are intended to obstruct (or at least reduce) the mobilization of anti-government sentiments. Where governments are already threatened by organized groups, such pre-emptive shutdowns are likely to undermine opposition preparations for collective attacks (see also Herreros and Criado, 2009), and should therefore occur *prior* to major episodes of state violence aimed at deterring further rebellion. Whereas sustained outages might be an effective tool to impede the long-term mobilization of opposition groups, they run the risk of motivating dissatisfied people to join protests against this extreme form of censorship.

In contrast to such widespread and enduring content censoring, empirical evidence shows that most incidences of actual 'blackouts' cover a relatively short timespan, not least because long-term outages affect a country's reputation and economic capacity (Howard, Agarwal and Hussain, 2011: 220). Since short periods of denied access to the internet are not likely to affect sustained opposition activity, it is plausible to assume that complete shutdowns are implemented in anticipation of gaining temporary advantages over a violently resisting opposition. Short-term shutdowns are therefore likely to be part of a repressive response towards an already mobilized opposition, in order to impede their capability to successfully implement, as well as the ability to coordinate, larger attacks against the state. The following sections discuss in more detail the incentives and costs associated with temporarily shutting down the internet.

Incentives for temporary network outages

The first anticipated benefit of temporarily shutting down network services is related to opposition groups' capabilities to effectively carry out attacks against the military. As discussed in chapter 3, highly sophisticated location-based services, such as Google Earth and Google Maps are used by ill-equipped rebel groups to locate military targets, and to calibrate weapons accordingly (Miller, 2012; Keating, 2013). Faced with an army that is superiorly equipped with weapons, technology, and trained soldiers, opposition groups frequently conduct asymmetric or 'irregular' warfare (Kalyvas and Balcells, 2010), where reliance on all available means of combat is pivotal.

Secondly, the ability to coordinate personnel, material and last-minute strategies via mobile phones and the internet is a vital channel by which opposition groups are able to organize attacks and resistance against the government. The presence of virtual communication channels, accelerated by smart phones, has increased the value of disseminating information and using it as a 'coordinating force [...] dramatically' (Shirky, 2008: 159). Additionally, '[b]logging, tweeting, podcasting, and taking pictures and videos and uploading them to Flickr and YouTube can all be done at near-zero financial cost (MacKinnon, 2012: 24)', making these tools available to anyone with a working internet connection. Recent studies report that social networking may in fact lead to increased participation in protest (Tufekci and Wilson, 2012), although these processes are less likely to be permanently affected by short intermissions in access. The internet provides a channel of communication that is fast, cheap, and harder to manipulate than more traditional, centralized media types such as the radio or newspapers (see Edmond, 2011: 25). A recent report interviewing members of the Free Syrian Army (FSA) supports these findings, reporting that

> '[e]very fighter seems to have at least one mobile phone, used to speak with families, Skype [...], and even advise Syrian soldiers how to defect to the opposition. Some note the difference a generation can make to the fate of their challenge against the government – and providing video evidence of atrocities and war crimes that are corroding the legitimacy of the regime.' (Peterson, 2012).

The threat posed by the increased abilities to coordinate, disseminate information, and even incite military soldiers to join the rebellion is likely to be immediately apparent to authoritarian rulers who fear for their political survival. In response, in a campaign aimed at repressing and possibly even eliminating the opposition, shutting down these communication channels can constitute a rational policy decision.

Where state-run mobile phone and internet services are generally accessible, opposition groups are likely to make use of them. Longer periods without access to state-provided network services should increase the probability of rebel groups finding alternative means and services, such as satellite phones and modems, network access via neighbouring countries, or dial-up connections (when landlines are accessible).[1]

[1] Opposition groups might decide to reorganize entirely and banish mobile and virtual communication from their coordination repertoire. In such instances, the anticipated benefits from shutting down network services are likely to be low.

The portfolio of surveillance and censorship methods to which state-run connections can be subjected is diverse. However, compared to using alternatives such as satellite networks, reliance on tools used by the majority of a country's population might still provide more security than using less broadly used channels such as satellite connections. New research on IT security demonstrates how satellite mobile devices produce traceable signals that allow governments to simply locate users and trace messages (see Driessen et al., 2012).[2] Locating and targeting rebels who are communicating outside of conventional structures offers a clear coercive advantage for the state when compared to their use of ordinary network connections.[3] Consequently, although opposition groups might possess alternative ways of connecting via cell phone and internet, the increased usage of satellite devices is likely to improve the regime's capability of identifying armed fighters among the civilian population.

The costs of network outages

Shutting down network services not only affects the opposition and its supporters, it in fact affects a country's entire population, not least due to losses in economic revenues. As discussed in section 3.2, when Egypt shut down its network access for only five days in 2011, the country's economy suffered at least $90 million loss in revenues (Howard, Agarwal and Hussain, 2011: 231).

Above all, as the theoretical argument put forward in this dissertation stresses, governments pursuing a counterinsurgency strategy in response to political threats are highly dependent on information provided – willingly, or unwillingly – by the civilian population (Lyall, 2010). Cell phone and internet access considerably facilitate communication for civilians willing and able to share crucial information on the location and activities of opposition fighters, without said fighters noticing the correspondence. This "'human intelligence" mechanism (Shapiro and Weidmann, forthcoming: 5)', should ultimately provide the state with an advantage over the rebellion. Furthermore, the Syrian regime has a proven history of extensive and invasive internet surveillance, which has allowed it to locate and target those deemed a threat on many occasions (OpenNet Initiative, 2009). Internet surveillance, however, only works, where the network accessibility is provided.

[2]Despite research in this field being comparatively new, news reports quoting researchers offer support:

> 'Radio direction finding and signals intelligence could easily be deployed in this scenario to figure out where the opposition is communicating from,' said John Scott-Railton, a research fellow at the Citizen Lab, an organization at the University of Toronto that focuses on Internet security (Perlroth, 2013).

Security researcher Jacob Appelbaum contends:

> Satellite phone systems and satellite networks are unsafe to use if location privacy or privacy for the content of communications is desired. These phone protocols are intentionally insecure and tracking people is sometimes considered a feature. (Jacob Appelbaum, quoted in York and Timm, 2012).

[3]The killings of two journalists in Homs in February 2012 support the notion that governments are making use of this technology. Security specialists contend that the Syrian government is likely to have directly targeted the houses from which they had traced the phones' signals (York and Timm, 2012).

Given the incentives and costs for governments in shutting down their networks, cutting all internet access is likely to be most effective in stifling opposition capability when used on an infrequent, temporary basis. However, 'overuse' of this most extreme form of censorship is likely to be counterproductive: If network disruptions precede all forms of military actions and occur on a regular basis, opposition groups will be able to use them as an 'early-warning system'. Following on from this, I argue that network outages are likely to be consistently associated with increased fighting. Conversely, not all periods of intense fighting are likely to be accompanied by a shut down networks. In short, disruptions will consistently be part of larger military campaigns, whereas not all military campaigns will entail disruptions.

Coercive response or cover-up?

Contemporary conflicts are being documented and simultaneously shared with the outside world through the help of the internet (Diamond, 2010). An alternative explanation for outages in contentious situations could be the government's intention of covering up and hiding violent acts from international scrutiny. Where a regime already receives increased international attention for repressing its citizens, cutting network activities might be part of an attempt to limit the extent of information leaving the country. Disruptions could present a chance to commit more large-scale acts of violence against the population, attempting to 'drain [...] the sea' (Valentino, Huth and Balch-Lindsay, 2004: 385) and eliminating the opposition, without creating a national and international audience, thereby potentially increasing the risk of sanctions, interventions, or even a referral to the International Criminal Court. Although autocratic regimes frequently engage in large-scale violence even when the international community is watching, the less real-time information is available, the more likely leaders will be able to plausibly deny responsibility for these atrocities (Mitchell, 2004).[4] The unprecedented number of journalists being killed in Syria demonstrates that the regime is evidently not indifferent to coverage of the conflict. According to the Committee to Protect Journalists (CPJ), Syria was the most dangerous country to be working in as a journalist in 2012 and 2013, with at least 61 journalists killed between 2011 and 2013 (Beiser, 2013). CPJ further reports on at least 60 kidnappings of press staff in 2013, as well as of journalists being tortured to death (ibid).

The cover-up argument has been voiced by international advocacy groups, such as Amnesty International, who has stated that:

> '[a]s fighting intensifies [...] we are extremely worried that the news that internet and mobile phone services appear to have been cut throughout Syria may herald the intention of the Syrian authorities to shield the truth of what is happening in the country from the outside world'.[5]

The intended effect of a disruption should therefore be an 'unobserved' increase in government repression. Although more atrocities are occurring, the groups

[4] For coverage on the Syrian Regime's plausible deniability of other events, such as the chemical attack in Ghouta in 2013, see Beaumont (2013).

[5] Ann Harrison, Middle East and North Africa Programme, Deputy Director Amnesty USA (Amnesty International, 2012).

collecting and disseminating the details on these events might have reduced access to their informants who usually provide evidence on individual victims.

The victims of violence during outages

Whether the victims of government violence change during internet outages likely depends on the government's motivation for the shut-down. Following international law, the literature on state repression broadly differentiates between combatant and non-combatant victims, while acknowledging that this distinction is oftentimes intentionally or unintentionally ignored by governments (Downes, 2008). Threatened governments are likely to intentionally conflate the status of combatants and civilians in irregular civil conflicts, where the organization of the opposition is opaque and front-lines between groups are unclear (Valentino, Huth and Balch-Lindsay, 2004; Balcells, 2010). As such, the Syrian regime has conducted the type of atrocious campaign against both rebel fighters and non-combatants that assumes anyone not showing explicit support is opposed to them.

If the intent behind shutting down the internet is to cover up prosecutable war crimes against unarmed civilians, it is plausible to assume that the composition of victims of government violence changes during outages, since the explicit focus of these disruptions would then be to attack as many civilians as possible. Empirically, a significant increase of - possibly unobserved - violence that only occurs *during* outages should therefore indicate a higher proportion of civilians killed.

With recent research indicating that censoring or blocking of the internet can lead to an increased turnout of protesters taking to the streets, an alternative reason for a higher proportion of civilian casualties would be if governments decided to then violently crack down on these protesters. Empirically, a substantial increase in violence in the *immediate aftermath* of outages would offer support for this scenario.

In this chapter, I argue that governments selectively implement outages as part of particularly repressive responses to increased armed opposition resistance. Where shutdowns are part of a concerted repressive response, the most substantial increase in violence should begin *prior* to, and then continue throughout, outages. In the midst of fighting, it is plausible to assume that anyone deemed as belonging to the opposition - including armed fighters and civilians standing by - are likely to be indiscriminately attacked.

7.2.1 Testable implications

If governments use network disruptions as a military tactic that forms part of a concerted repressive offensive against opposition groups, a main observable outcome is an increase in the activity of pro-government fighters during and in the immediate time surrounding outages. Pro-government fighter activity is measured by the number of people killed by the regime. According to the theoretical expectations laid out above, I expect short, unexpected network outages to be accompanied by significantly higher levels of military activity, and thus significantly higher numbers of people killed. The number of *actual* people killed is defined as the combined number of *documented* and *undocumented* fatalities. In order to understand whether disruptions are linked to higher

Table 7.1. Expected effects for network disruptions and violence

Empirical expectation	Documented violence	% Undocumented violence	Timing of increase in violence
Coercive Response	increase	no change	**prior**/ during disruption
Cover-up	no change / increase	increase	**only during** disruption
No effect	no change	no change	

levels of violence, it is crucial to account for changes in documented and undocumented violence, since changes in communication technology might have an effect on the documentation process. Actual levels of violence, meaning documented and undocumented cases combined, are not directly observable. I use the estimation technique presented in Chapter 6 to predict the number of undocumented cases. The main empirical expectation, given that disruptions are part of a state's set of military tactics is:

- **Empirical Expectation:** All else equal, periods of network disruption are accompanied by a significant increase in actual conflict fatalities.

The main alternative explanation is that governments use disruptions to cover-up their atrocities:

- **Alternative Expectation:** All else equal, the proportion of undocumented conflict fatalities increases significantly during network disruptions.

Whereas the number of undocumented cases is seldom zero, the alternative explanation for why governments cut their networks is that they do this to cover up their crimes, which means the dark figure of unreported cases should increase disproportionately to the number of documented cases. A further possible scenario is that governments intend to cover their tracks, but that they are unsuccessful at doing so. An additional factor is therefore considered, which is the timing of violence versus network disruptions. If governments care about the news of atrocities travelling beyond the battle grounds, they are likely to only *commence* with the violence once the network is disconnected. Starting a campaign of violence and then shutting out the international community is likely to raise more awareness than before. In short, cover-up campaigns should show no signs of an increase of violence *prior* to the outage, and if successful, should hide a large increase in undocumented fatalities *during* the disruption. In contrast, increases in violence immediately preceding disruptions are consistent with the coercive response hypothesis. I expect increases in military activity *prior* to and *during* disruptions to indicate the strategic value of shutdowns in government repression policy. Table 7.1 summarizes the expectation of the main hypotheses and the alternative explanation for documented violence, the percentage of undocumented violence, as well as the timing of violence prior to and on days with network disruptions.

7.3 Data and empirical strategy

Network outages in the Syrian civil war

Syria's government has a demonstrated history in blocking content on the internet (OpenNet Initiative, 2009; Deibert, 2008). Since the start of the civil conflict, there have been two main types of internet disruptions: National, large-scale outages, and smaller regional variations in accessibility. This chapter only analyzes large-scale national incidences of complete network outages. Evidence on the trajectory of these outages suggests that technical failures as the possible cause can be ruled out (Gallagher, 2012). These outages have occurred at irregular intervals, without being anticipated by either the international media or the opposition groups.

Local intermissions generally occur in parts of the country that are already controlled by opposition groups, most notably the Northern governorates Ar-Raqqah and Al-Hasaka. These two governorates have experienced limits in accessibility to the internet for most of the period under investigation. Syrian security experts contend that these deteriorated connections occur in regions where the opposition has taken control of territories, in an effort to withhold public goods from a population that is 'collaborating' with the regime's enemies.[6]

The country-wide outages for the period between March 2011 and September 2013 are determined through the information collected by the Google Transparency reports on traffic disruptions in Syria since March 2011.[7] Suspensions of traffic that lasted between a few hours and three days occurred in June 2011, July 2012, November 2012, January 2013 and twice in May 2013.[8] To account for the empirical expectations of the theory, I include three different treatments for network outages. The first dichotomous variable takes on the value of 1 on days where the traffic was disrupted, and a 0 for days of normal connection. The second variables sets the treatment at $t-1$, the day *prior* to the disruption. The third variable looks at the time window of the disruption, and codes the day prior to, the days of the disruption, and the following day as 1, and the rest as 0. To control for decreasing or increasing effects over time, I include a measure that accounts for the number of previous outages, as well as a variable that measures the time since the last outage, as recent outages might positively or negatively affect the dynamics of violence.

Documented conflict fatalities

Table 7.2 provides an overview of the documented daily fatality counts by governorate. The highest number of fatalities are reported in Rural Damascus, Homs, Aleppo and Idlib. Rural Damascus also witnessed the maximum number of fatalities per day for the period from March 2011 to September 2013, which doubles the number of any other governorate. The outer governorates of Tartus, Ar-Raqqah, Al-Hasaka, Quneitra, and As-Suwayda all have comparatively low

[6] Personal communication with Dlshad Othman (Kurdish Syrian Activist and Internet Freedom Fellow), Anas Qtish (Syrian Blogger, Electronic Frontier Foundation), and staff of the Syrian Digital Security Monitor (`https://syria.secdev.com/`).

[7] `http://www.google.com/transparencyreport/traffic/`

[8] The fraction of normalized worldwide traffic in Syria is presented for a sample of outages in the online appendix, Figures 2, 3, and 4.

Table 7.2. Summary statistics, documented fatality counts

Governorate	Min	Max	Mean	St. Dev.	$\sum$
Rural Damascus	0	645	24	34	22,155
Homs	0	265	19	20	17,823
Aleppo	0	255	16	19	14,665
Idlib	0	164	13	14	11,831
Daraa	0	185	11	13	9,898
Hama	0	168	10	13	9,100
Damascus	0	137	8	12	7,622
Deir ez-Zor	0	117	7	10	6,472
Latakia	0	66	3	5	3,203
Tartus	0	125	3	6	2,413
Ar-Raqqah	0	29	2	3	1,458
Al-Hasaka	0	31	1	3	1,141
Quneitra	0	42	1	2	786
As-Suwayda	0	8	0	1	433

numbers of documented violence, further supporting evidence that these areas were not at the center of clashes between regime and opposition forces for most of the conflict period under investigation. In the absence of consistent struggles between the government and opposition, the expectation is that the effect of network outages is likely to be less pronounced than in governorates such as Rural Damascus, Aleppo, Homs and Idlib.

7.4 Analysis I: Network outages and documented killings

7.4.1 Descriptive evidence

The descriptive difference of documented daily killings during network outages is presented in Figure 7.1, which maps the average difference in daily killings between days where the internet is turned on, and days where the country is disconnected. In the North-East of the country, opposition groups have established quasi-administrative structures (MacFarquhar and Saad, 2012), and consequently have been cut off from central government services, including the internet, which means that national outages are likely to display little effect in these regions. Evidently the effect depends on the degree of armed confrontation between opposition and government groups, which means that sub-national variations need to be taken into account in the analysis. The North-West of the country, where Aleppo and Idlib are located, show more than 20 additional fatalities on days where there is no internet across the country, compared to other days during the period under investigation. The conflict hotspots of Rural Damascus and Homs in the center of the country show an average daily increase of more then 30 fatalities.

To further investigate the relationship between violence and disruption it is useful to visually inspect the dynamics of violence and disruptions across time. Figure 7.2(a) plots the daily counts for Hama from April to August 2011, marking the disruption days in June in yellow. A sharp increase in violence on

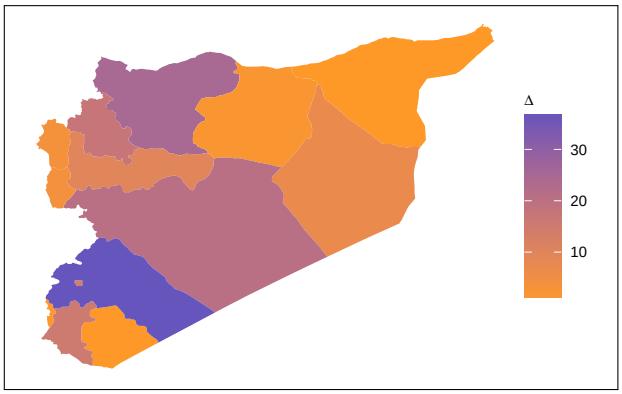


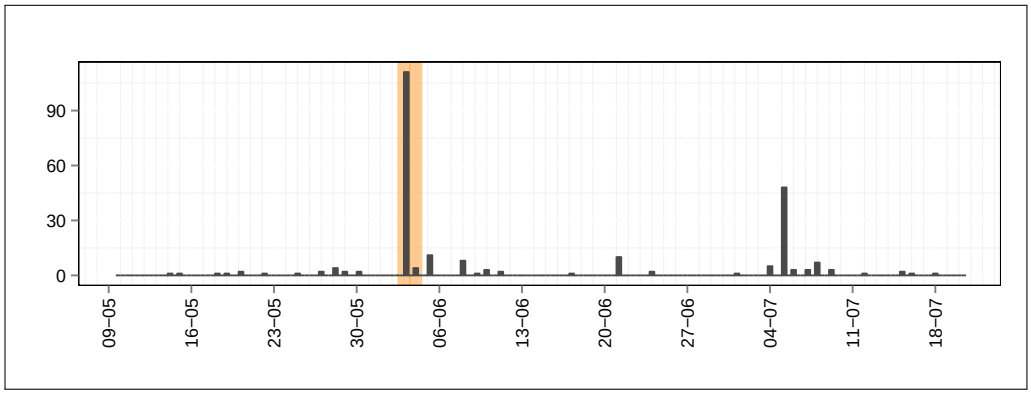
Figure 7.1. Mean difference in daily killings between days with and without internet, Syria, March 2011 - September 2013.

the first day of the outage is clearly visible in this graph. A different trend is shown in Figure 7.2(b), which plots the daily counts for Rural Damascus across May to August 2012. As can be seen quite clearly, the number of killings rises substantially on the day *before* the blackout, decreases on the day concerned to a still high number, and increases slightly on the following day. This visual inspection indicates that the association between disruptions and increases in violence moves beyond the mere outage days. As discussed above, network disruptions implemented as part of a coercive response need not necessarily be implemented prior to the commencement of fighting. Shutting down the network amidst fighting is likely to constitute a role in a military strategy.

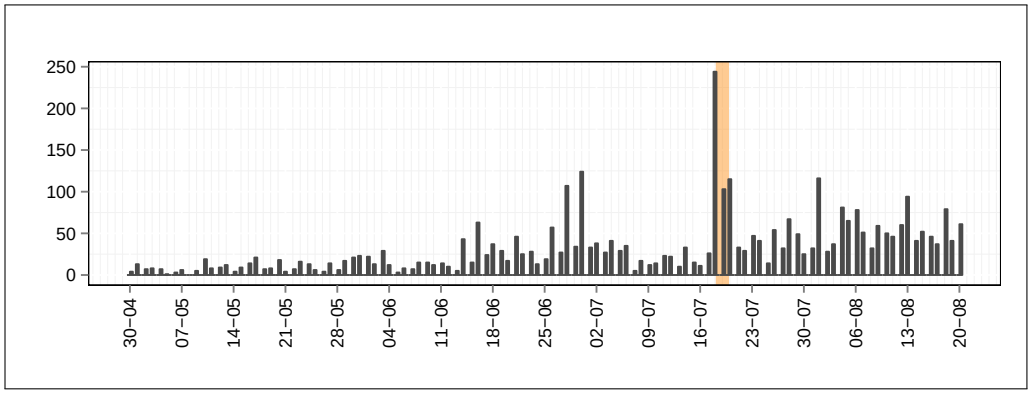
To further investigate the difference in violence with regard to the theoretical predictions, Figure 7.3 plots simple difference of mean tests at the national and the governorate-level for different 'treatments', by comparing all other days with the day prior to $(t-1)$, the day of the outage (t), and the time window$(t-1, t, t+1)$ surrounding the disruptions. If the state is using violence in response to a dispersion of protests, we should see the largest difference on the day, and the day following the disruption. If the government plans to cover up all violence occurring during the disruption, we should either see *no* increase in killings during the disruption, or before or after the disruption, or we should *only* see an increase *during* the disruption. All tests show that the means are significantly different from zero, and the highest average difference at both levels of aggregation can be found at $t-1$, the day prior to the outage, offering further evidence for the coercive response hypothesis.

7.4.2 National-level evidence

Since the outages occur at a national level, the first step of the multivariate analysis examines the national effect with daily data from the 15th of March 2011 until the 30th of September 2013. In view of the fact that the conflict in Syria has intensified over time, the number of killings follows a generally increasing, non mean-reverting trend. To account for these dynamics, I estimate a Poisson exponentially weighted moving average model (PEWMA), as formalized by Brandt et al. (2000). The PEWMA is a structural time series model that nests a Poisson model, where observed counts at time t are modeled as a



(a) Hama 2011



(b) Rural Damascus 2012

Figure 7.2. Violence and network disruptions, Hama 2011 & Rural Damascus 2012.

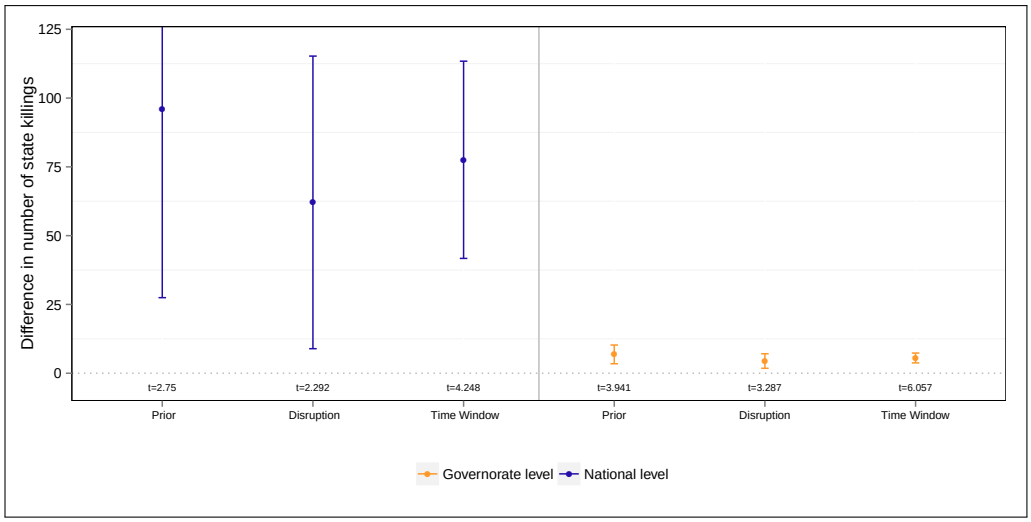


Figure 7.3. Difference of means tests for days with and without disruptions.

weighted average of counts at previous time points (Brandt et al., 2000: 827).[9]

[9] The model estimates a hyperparameter ω that accounts for dependence in the event counts over time, where values close to 0 indicate more dependence, and values approaching 1 indicate

Table 7.3. National-level time-series model: Disruptions and violence

	Model 1	Model 2	Model 3
Pre disruption	0.233		
	(0.105)		
	26.3%		
Disruption		0.083	
		(0.098)	
		8.6%	
Time window			0.383
			(0.087)
			46.7%
Last disruption	0.001	0.001	0.001
	(0.000)	(0.000)	(0.000)
ω	0.063	0.063	0.064
	(0.002)	(0.002)	(0.002)
N	851	851	851
LLF	-4,581.051	-4,582.848	-4,573.671
AIC	9,166.101	9,169.696	9,151.343

Poisson Exponentially Moving Average (PEWMA) Model.

Standard errors in parentheses.

Predicted percentage changes in italics.

Since the interpretation of the coefficients is not straightforward, I calculate the predicted percentage changes in killings for the same three treatments used in the difference of mean tests. Table 7.3 reports the results of the three models. The model predicts that on average, the level of violence increases by 26.3% on the day prior to an internet outage. During the actual blackout, violence is predicted to increase by 8.6%, and when looking at the time window, the average increase is predicted to be almost 47%. Although not all governorates are affected by fighting in the same way, the aggregate national evidence offers further support for the hypothesis that outages are preceded and accompanied by significant increases in violence.

7.4.3 Regional evidence

Since the level of contestation varies substantially across different regions in Syria, the degree to which states increase their offensives in conjunction with the outages is likely to differ from region to region. To obtain estimates for each region in Syria, I estimate a time-series cross section fixed-effects poisson model, where the 14 governorates are the fixed units. I simulate the expected change in the number of regime fatalities in each Syrian governorate between a day where all networks are available and a day where they are shut down. Figure 7.4 shows all 14 governorates along the x-axis and plots the expected change in

few dynamics, and a data structure that could potentially be modeled with a conventional Poisson model. I thank Patrick Brandt for providing the estimation code on his website: `http://www.utdallas.edu/~pbrandt/pests/pests.htm`.

fatalities including the 95% confidence interval on the y-axis. The yellow lines show the expected change on the day of network outages, and the blue lines show the expected effect on the day prior to an outage. None of the confidence intervals include zero, which means that the days with network disruptions witness statistically significant higher levels of violence across all Syrian governorates. In both models, the substantive effect varies clearly across governorates, which is not surprising given the significant differences in the levels of violence experienced. In Homs and Rural Damascus, days without internet (yellow lines) experience at least three additional incidences of lethal violence when compared to days with regular access. The effect of network outages on violence seems less pronounced in peripheral regions such as As-Suwayda, Al-Hasaka and Quneitra.

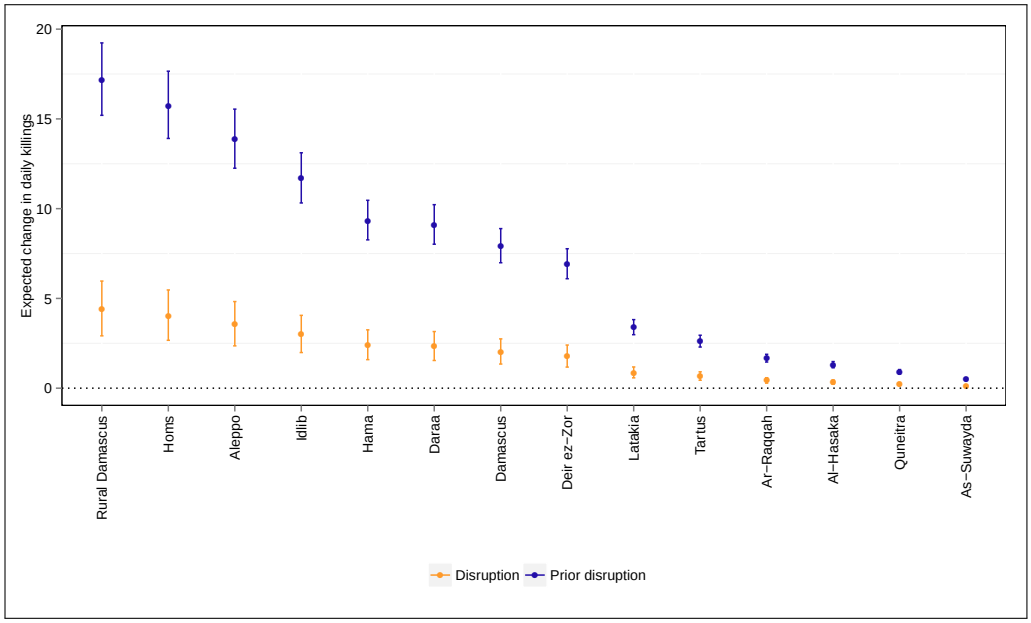


Figure 7.4. Expected change in daily killings, given network disruptions.

On days prior to internet disruptions, the estimated expected changes in violence are substantially larger. The expected increase in conflict fatalities in Homs, Rural Damascus, Aleppo and Idlib is above ten killings. Both models offer support for the coercive response argument: Governorates in Syria experience a significant increase in conflict deaths perpetrated by the regime on days where the regime shuts down network services. Furthermore, a first substantial increase in violence occurs one day prior, an increase we would not expect if the regime were interested in covering up atrocities during blackouts, or cracking down on a higher number of protesters as a result of the outages.

Due to the dynamic nature of conflict violence, it is important to test whether the results are being driven by general conflict trends in the data. In addition, I test whether the results hold when replacing absolute levels of documented violence with the first difference, where the dependent variable only reports the change in fatalities between two days. The first model in Table 7.4 tests for an increase in violence on the days prior to disruptions, the second model tests for the first day of actual disruptions, and the last model tests for the day afterwards,

Table 7.4. First difference model: Network disruptions and changes in violence

	Model 4	Model 5	Model 6
Intercept	−0.270	−0.251	−0.256
	(0.187)	(0.187)	(0.188)
Pre disruption	5.743		
	(1.596)		
First day		3.539	
		(3.323)	
Post disruption			1.044
			(1.466)
Last disruption	0.002	0.002	0.002
	(0.001)	(0.001)	(0.001)
# Disruptions	0.384	−0.437	0.381
	(0.376)	(0.852)	(0.376)
Diff_{t-1}	−0.599	−0.600	−0.600
	(0.009)	(0.009)	(0.009)
Diff_{t-2}	−0.273	−0.273	−0.273
	(0.009)	(0.009)	(0.009)
R^2	0.281	0.280	0.280
Adj. R^2	0.280	0.280	0.280
N	11,914	11,914	11,914

in order to investigate whether violence continues to rise. As expected, the most statistically significant and substantive increase is found on the day prior to the disruption. The effect on the actual days with outages is not as pronounced, and there seems to be no enduring effect once networks are turned back on again.

7.4.4 Placebo Tests

Given the small number of 'treatment' incidences over the time period of three years it is important to check whether these results might be due to chance.[10] A useful way to do this is to use a placebo test (see Dafoe and Tunón, 2014). In order to maintain the structure of the treatment variable, I create a series of time-shifted placebos (see Reynolds, 2014; Dube, Kaplan and Naidu, 2011), where the treatment is moved between $t - 30$ and $t + 30$ intervals, to account for the time period one month prior to and one month following each outage.[11] For each placebo, the national-level time-series count model is estimated, and the predicted percentage change of violence on days with the treatment saved. Figure 7.5 plots the predicted changes, as well as the actual treatments at t (the disruption) and $t - 1$ (the day prior the disruption). The majority of all placebos predict a change in violence that is negative, or less than +2%. Importantly, the predicted change in the immediate aftermath of the outage are either zero, or

[10] Recent revelations by NSA whistleblower Edward Snowden suggest that the November 2012 outage might have occurred due to a technical failure brought about by the NSA (Bamford, 2014).

[11] The placebos created were at t [-30,-25,-20,-15,-10-,-5,+1,+5,+10,+15,+20,+25,+30].

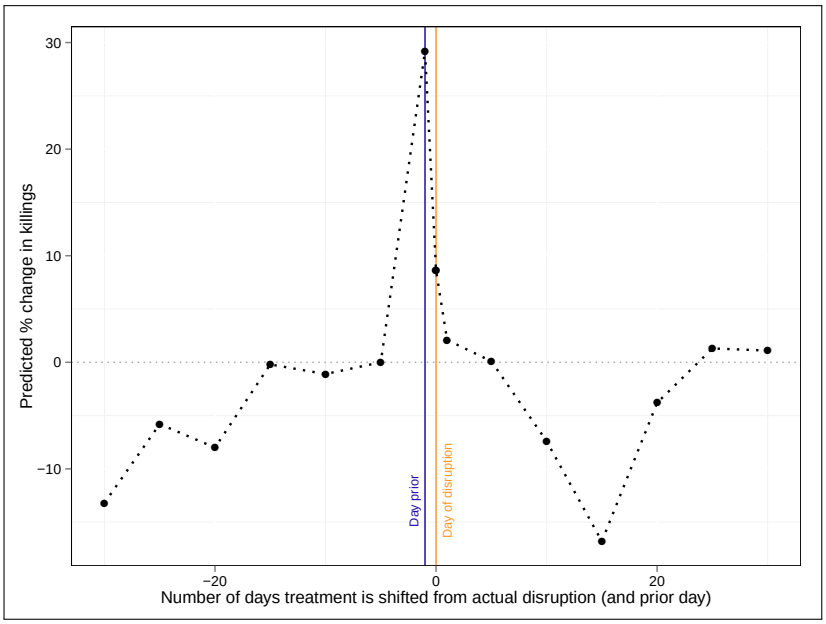


Figure 7.5. Time-shifted placebo treatment test.

negative, further confirming that governments are not responding to an increase in protest during the disruption.

7.5 Analysis II: Network outages and documentation patterns of violence

The conflict in the Syrian Arab Republic has led to one of the most sophisticated real-time documentation efforts in the history of casualty recording with countless groups and organizations working to keep track on the violence. As in all conflicts, however, it is impossible to determine the true population of conflict fatalities via documented data. Human rights groups are doing their very best to document all violence that is documentable, but for analyses such as the one attempted in this study, it is of paramount importance to obtain an estimate of *all* fatalities, not just of those documented. Studies examining the effect of information technology on the intensity of violence are particularly sensitive to potential biases in conflict data that might arise precisely because of changes in said technology. This study is no different, and the cover-up hypothesis, which assumes that the outages were implemented to cover-up state-led atrocities, even supports this claim.

7.5.1 Variation in reporting before and during disruptions

One way to test for the cover-up hypothesis is to determine whether the level of underreporting differed substantially on days without internet, compared to days with network access. Since four datasets are de-duplicated and matched for the entire observation period, the overlap structures of fatalities that were recorded by 1, 2, 3 or 4 sources can be used to estimate the number of fatalities that were not documented by any source, as described in Chapter 6. Log-linear

capture-recapture estimation follows this simple intuition and has been used to estimate fatalities in a multitude of conflicts (see Lum, Price and Banks, 2013).[12] I isolate the number of documented fatalities by governorate for the days without internet, and estimate the number of undocumented killings for each of these periods and regions separately. Since the degree of underreporting is likely to vary across time and space, I select the fixed period of a week prior to each network disruption and estimate the level of underreporting at the national level, and for each respective governorate as well. I then compare the levels of regional underreporting for the immediate time period prior to the disruption with the underreporting during the disruption, in order to assess whether or not disconnected days lead to systematic underreporting of violence.

For example, during the internet blackout on the 3rd and 4th of June 2011, 24 victims were documented in Aleppo. Only 7 of these victims are reported in all lists, the remaining victims were reported by a combination of less than all sources. Capture-recapture estimation reveals that it is highly likely that 42 individuals were killed in this period (with a 95% confidence interval of [25, 150]), which means that 42.7 % of all victims went undocumented in those two days. In the week prior to the June outage, 133 victims of regime violence were documented in Aleppo, of which only 46 were known by all sources. The estimated number of actual regime fatalities is 167 [c.i.: 146; 212], which means that 20.4% of all cases went undocumented.

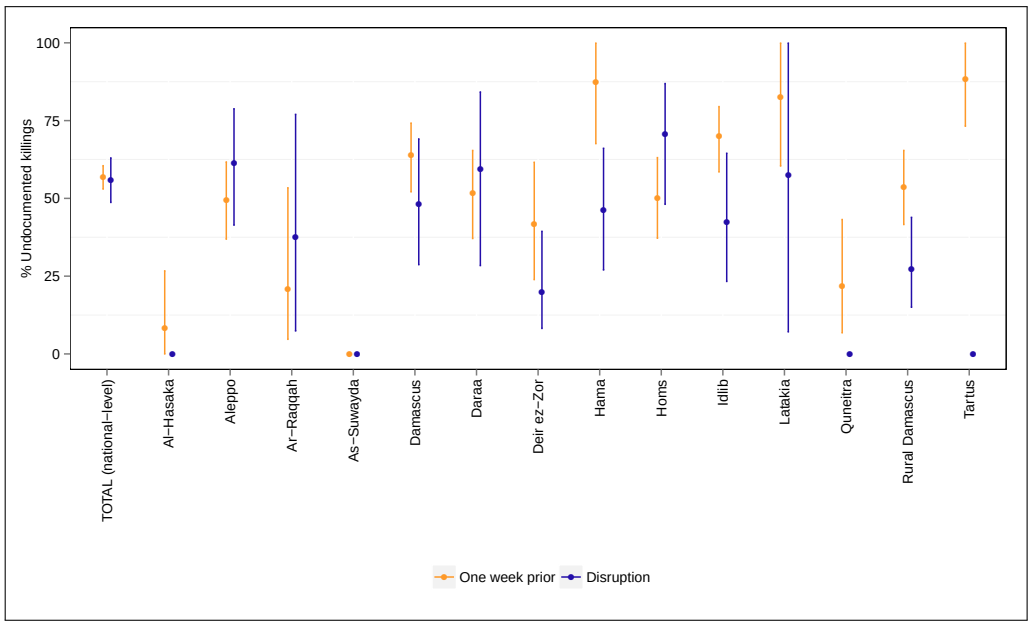


Figure 7.6. Per cent of undocumented fatalities (of actual number) one week prior to, and during disruptions, by governorate.

Figure 7.6 shows the level of underreporting for the week prior to, and during internet outages at the national and governorate-level, including 95 %

[12]Log-linear poisson models for capture-recapture estimation are implemented in the R package Rcapture (Baillargeon and Rivest, 2007). Log-linear models are effective in dealing with capture heterogeneity and list dependencies, two of the main challenges when estimating conflict fatalities (see Manrique-Vallier, Price and Gohdes, 2013).

confidence intervals. At the national level, the degree of underreporting is almost exactly the same, and at the governorate-level, only Hama displays a significant difference, but in this case underreporting was significantly higher *prior* to the outage, which poses no problem for the validity of the results. The results suggest that documentation patterns are not systematically linked to network outages. Whereas variation in reporting across governorates is partly visible, it is likely to be driven by other factors not addressed in this study. I also estimate the level of underreporting for each outage separately, the results of which are reported in Figure 7.7.

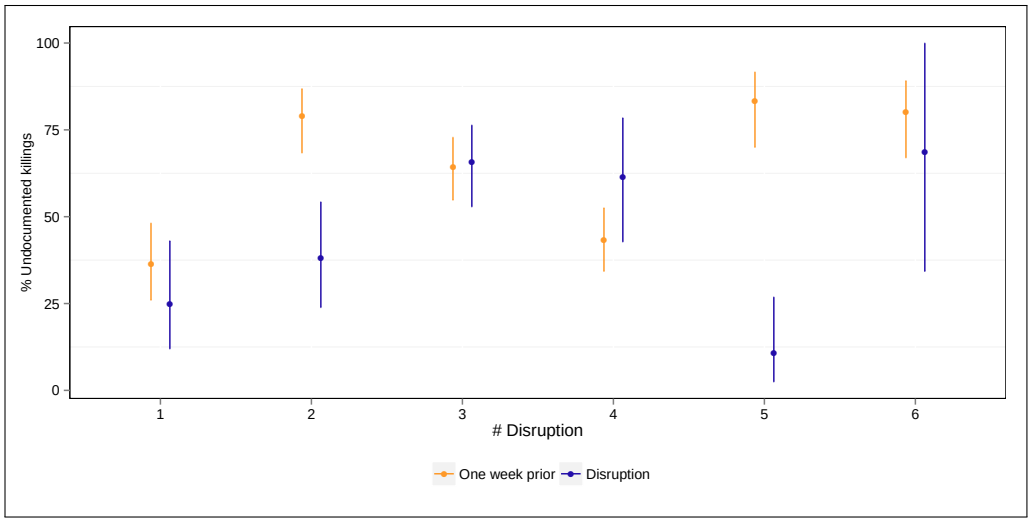


Figure 7.7. Per cent of undocumented fatalities one week prior to, and during disruptions, measured for each disruption separately.

7.6 Summary

Censorship of the internet is nothing new: authoritarian regimes intent on maintaining the status quo within their country have been relatively successful at manipulating content in their favor (Morozov, 2012; Rød and Weidmann, forthcoming). What has remained unclear to date, however, is to what extent extreme forms of censorship - such as the cutting of all connections - have the potential for constituting a tactic within larger military offensives. The results of the analysis of network outages and daily conflict fatalities in Syria suggest that regimes implement large-scale disruptions selectively and purposely in conjunction with launching larger battles. Evidently, not all battles are accompanied by outages, but when they are, they tend to be preceded by a substantial increase in violence.

Even in conflicts that are under as much national and international scrutiny as the current case of Syria, it is important to analytically distinguish between the empirical implications for *documented* violence, and the empirical implications for *actual* levels of violence: cases that are observed and those that are either intentionally or unintentionally hidden from documentation. The theoretical expectations advanced in this study clearly distinguish between implications for documentation of violence, and violence in general. Distinguishing between the

documented and the dark figure of violence improves the analytical leverage of the study's findings: the fact that undocumented violence in Syria is not systematically affected by short disruptions offers important support for the coercive response hypothesis.[13] The comparison and testing of these competing theories (military strategy vs. cover-up) would not have been possible without an estimation of the unreported number of state killings. The dependency of the observed data on information accessibility would not have been known. Estimating the degree of underreporting, however, also demonstrates the variability in documentation. Cases where more violence is hidden from view during disruptions might turn out to be a welcome side-effect for governments seeking to maintain international legitimacy and internal control.

This chapter has attempted to understand why governments might have an incentive to include the disruption of internet and cellphone service in their military strategy. I have argued that the scarce and sudden disconnection from essential communication networks is likely to weaken opposition groups' propensity to organise, but further research is needed in order to understand whether this is in fact the case, and if so, what the exact underlying mechanisms are that allow network failures to get in the way of effective information dissemination. This analysis has analysed the effects of nation-wide outages. The next chapter investigates how regional variations in the accessibility of the internet affect strategies of violent state repression.

[13] Incidences where the shutdown lasts much longer might produce very different results.

8

Information, connectivity, and strategic coercion

8.1 Introduction

This chapter investigates how a government's ability to censor and limit the flow of information feeds into its choice of coercive tactics. I analyse how sub-national variations in internet accessibility in Syria affect the government's use of repressive strategies. The Syrian regime has made use of a wide array on surveillance measures to spy on its citizens, in particular all those it has deemed threatening to its survival (see Galperin and Marquis-Boire, 2012; Galperin, Marquis-Boire and Scott-Railton, 2013). Likewise, it has, at different points in time, limited the accessibility of the internet across the entire country, and in individual regions significantly. I commence with a brief overview of research dealing with the effect of information control on coercion, and then theorize how the trade-off between censorship and surveillance affects a government's ability to effectively repress those it deems threatening to its political stability. The conditions under which surveillance will be more effective, as well as conditions under which censorship will be more effective are discussed, and empirical expectations for each condition formulated. The empirical part of this chapter presents the survey data used to assess the degree of censorship in each Syrian governorate between June 2013 and April 2014, and discusses the steps involved in obtaining an accurate measure of targeted and untargeted repressive tactics used by the Syrian government across the same period of time.

I use supervised machine-learning to analyze over 60,000 records of killings perpetrated by the Syrian Regime in the ongoing conflict, and classify them according to their event circumstances, to arrive at a categorization between targeted and untargeted acts of repression. To account for variations in the level of reported violence, I use capture-recapture methods to estimate the total number of killings in both categories. This new indicator of strategic repression is used in a beta-binomial regression to estimate the effect of varying levels of censorship on the proportion of targeted (vs. untargeted) killings conducted

by the government at a certain time and place. I find that higher levels of information accessibility are consistently linked to an increase in the proportion of targeted repression, whereas areas with little or no access witness more indiscriminate campaigns of violence.

8.2 Information control and coercion

The relationship between violent state coercion and domestic dissent is both theoretically and empirically well-established (Moore, 1998; Davenport, 2007*a*; Conrad and Moore, 2010; Carey, 2010). Rulers weigh the perceived threat posed by protesting citizens and/or opposition groups against their perceived strength, and the likelihood of employing coercive policies increases as the imbalance between strength and threat grows (Poe, 2004). State repression can take on different manifestations in terms of scale, scope, and technical sophistication, but the forms most generally identified with coercion are physical integrity rights violations such as torture, imprisonment, killings, and disappearances.

The literature broadly distinguishes between targeted forms of repression, which are directed against dissidents and opposition members, and repression that is directed against large parts or the majority of the population (Wood and Gibney, 2010; Krüger and Davenport, 2014). The distinction is not only theoretically important, as the effect of repressive actions is likely to vary depending on the type and scope of violence directed against the population (Kalyvas, 2006; Lyall, 2009; Zhukov, 2013). Faced with disgruntled citizens, governments have to choose the level and scope of coercive measures to be used: whereas too light a response against a few individuals might be ineffective at obstructing serious attempts to challenge authority, too much violence might backfire by motivating previously agnostic or uninterested bystanders to join the protests or resistance movement. To maintain political control, rulers are likely to attempt targeting those individuals deemed threatening to the status quo. Where governments have identified large parts of their domestic population as a threat, repression is likely to be wide-spread and indiscriminate (Valentino, Huth and Balch-Lindsay, 2004).

For state repression to effectively stabilize the position of the ruler, the strategic use of censorship is critical (Frantz and Kendall-Taylor, 2014). Traditionally, more extreme forms of censoring information have been implemented by banning newspaper, radio and television stations and targeting journalists. Less extreme forms have included the surveillance of news agencies and banning and alteration of individual media content. In contrast to coercion, which is seldom directed at all citizens, the intentional restriction and altering of information flows indiscriminately affects an entire population (Wintrobe, 1998). As with indiscriminate violence, overly zealous censorship has the potential for backfiring, where citizens rate the absence of reporting as 'bad news' (Shadmehr and Bernhardt, 2013: 26), and thus lower their support for the ruling elite. Conversely, leaders fear that free and public criticism of their policies might jeopardize their standing even more (Kim, Whitten-Woodring and James, 2014).

The rise of citizen journalists working independently of traditional news agencies and providing content via internet platforms, and the sheer increase of information generated and shared across social networks have changed the repercussions and dynamics of censorship and rebellion considerably (Aday,

Farrell and Lynch, 2010). Communication via Skype, Facebook and Twitter, in particular accelerated by the availability of smart phones, has increased the value of disseminating information and using it as a 'coordinating force [...] dramatically' (Shirky, 2008: 159). Additionally, '[b]logging, tweeting, podcasting, and taking pictures and videos and uploading them to Flickr and YouTube can all be done at near-zero financial cost' (MacKinnon, 2012: 24), allowing ordinary citizens to provide documentation of state actions in situations where journalists have been expelled or arrested. Recent studies even report that social networking opportunities may in fact positively affect the political awareness of unsatisfied citizens (Reuter and Szakonyi, 2013), an attribute that is likely to increase the pool of potential recruits of anti-government rebellions (Tufekci and Wilson, 2012). Even when only looking at mobile phone connections, Pierskalla and Hollenbach (2013: 210) find that opposition groups strongly benefit from the availability of cheap communication tools, thereby increasing the incidences of organized, violent rebellions.

While the decentralized nature of the internet has enhanced the toolkits of activists and protesters, it has at the same time proved equally effective in broadening the repertoire of surveillance for governments (see Youmans and York, 2012), and forced leaders to adapt their censorship methods (Deibert, 2008, 2010; Roberts, 2014). Deibert and Rohozinski (2010) distinguish between three generations of internet control, where the first generation presents the most primitive form of blocking content, and the second and third generations involve more subtle ways of warantless surveillance and normative campaigns against critical information, intended to encourage self-censorship. With regard to first-generation controls (see also Howard, Agarwal and Hussain, 2011), the results presented in the previous chapter demonstrates that large-scale, short-term network outages in Syria were accompanied by substantive increases in government-directed violence, indicating that the shutdown might have been employed strategically to weaken the coordination capabilities of the opposition.

Looking at more sophisticated forms of censorship, King, Pan and Roberts (2013) demonstrate that the Chinese government has taken the mobilizing potential of the internet seriously, by showing that significantly more content inciting collective organization is censored than other types of content – even content that explicitly criticizes the ruling party. Both China and Russia provide fully supervised domestic 'cyberspaces' to their citizens, allowing them to engage relatively freely with each other online, and go to great lengths of only censoring what is – in their eyes – necessary to avoid internal unrests (Deibert and Rohozinski, 2010; King, Pan and Roberts, 2014). The combination of full surveillance on the one hand, and sophisticated censorship on the other hand has proven to be highly successful in both cases (MacKinnon, 2012). The effort made to maintain these systems speaks volumes about the trade-off modern rulers face. Allowing citizens to converse openly provides the ideal grounds for surveillance, but the proven potential for collective mobilization means there is risk involved as well.

Not all governments have sufficient resources to manually supervise and filter their domestic webspace, in the way China or Russia has done. Less expensive methods are revealed by recent research demonstrating the extent and ease of implementing warantless surveillance. Marczak et al. (2014) document extensive spyware and trojans used by the governments of Bahrain, Syria, and

the United Arab Emirates for 'eavesdropping, stealing information, and/or unmasking anonymous users' (Marczak et al., 2014: 1).

Obtaining information about the opposition does not necessarily require spyware: if a sufficient proportion of the population support the government, sharing knowledge on the location and planned activities of dissidents or insurgents is likely to occur via cell phones or anonymous tips online. Shapiro and Weidmann (forthcoming) analyze cell phone usage in Iraq and find that insurgent violence is significantly lower where increased mobile communication is available, which they argue is due to the civilian population sharing information with the government.

In conclusion, the majority of research intent on understanding censorship has therefore focused on how variations in media restriction might quell or instigate unrest. The informational value generated by the free exchange of information has remained somewhat neglected. Although the threat of collective action resulting from increased communication flows is evident, the exponential increase in user-generated content has provided important incentives for rulers to surveil their entire domestic population (Deibert, 2010; MacKinnon, 2012). The following section theorizes about how the control and limitation of information flows is likely to be linked to different strategies of government coercion.

8.3 Surveillance, censorship, and 'effective' repression

I assume that from the position of the government, a repression tactic is effective if it manages to eliminate or at least mitigate the threat posed, for example, by an insurgency, mass uprising or even smaller-scale protest. Ideally, such a tactic would involve identifying those individuals or organisations that are genuinely challenging the authority's position, as opposed to the neutral bystanders, and eliminating them, for example through arrest, expulsion, disappearance, or even violent death. To do this, leaders need identifying information (Kalyvas, 2006; Condra and Shapiro, 2012), and the opportunities for obtaining such information by surveilling online communication are immense.

Where citizens are able to access the internet and converse with others freely, they generate vast amounts of information that can be used by governments. Information gained via these services is used to create nuanced models of interaction, perceptions, location, intention, and network of collaborators for each citizen. Public and private events organized and distributed via social media, email, and other channels can easily be anticipated. Prospective participants of such events can be predicted and also placed under even closer surveillance. Each individual's *friends*, *followers*, call logs, newsletters, subscriptions and text messages can be used to obtain an understanding of how resistance movements are organized, and who constitutes the central actors. Once these particular 'threats' are identified, location-based services can aid in isolating and targeting them.

The use of surveillance to facilitate targeted arrests and elimination of threats to the political survival of regimes have long since been a part of the repertoire of coercive tools used by governments. The German *Staatssicherheit* was famous for its meticulous approach towards surveilling citizens in the German Democratic Republic, by listening in on phone calls, positioning staff in next-door homes, and getting neighbours, family and friends to spy on each other. Targeted arrests

based on the information gathered were then oftentimes conducted at night, making the disappearance of individuals less obtrusive.

In short, the free flow of information allows governments to effectively surveil citizens, and extract information needed to identify perceived central threats with relatively high levels of accuracy. The main tradeoff leaders face is, however, obvious: for surveillance of 'critical' information to work, critical information needs to be exchanged, which in turn can further strengthen those opposed to the central political authority. This fear led the Iranian government to limit access to the internet during the national elections in 2009. It also led the Sudanese government to disconnect its citizens from the internet in 2013 when protests sparked over fuel prices. In early 2014, it led the Turkish government to ban Twitter for two weeks. It led the Chinese government to block Instagram during recent mass protests in Hong Kong (Olesen, 2014). And it has led the Vietnamese government to continuously surveil and arrest bloggers.

Thus, in limiting the free flow of information, government are following the classic understanding of censorship: restricting criticism and calls for collective organization in order to maintain control and stability. This stability comes at the price of information loss (see also Lorentzen, 2013) and, if overused, loss of public support (Shadmehr and Bernhardt, 2013). I will now proceed to hypothesize under which conditions free (or freer) access to information, and under which conditions a more restrictive access to information will be more effective for governments intent on countering a domestic threat.

8.3.1 Incentives for surveillance

The ability to closely surveil the free flow of information will be particularly useful in situations where governments do not perceive the threat as coming from the overwhelming majority of the domestic population. In a first assessment, pending further information, leaders evaluate the extent of the challenge to their status quo. For example, if the source of unrest has already been traced back to a particular group – such as a student group, or a small ethnic minority – the surveillance benefits of permitting unrestricted information exchange will outweigh the costs. In this scenario, full surveillance allows governments to learn even more about their opponents, information that increases their chances of successfully putting an end to the anti-government activities. Equally important, it allows for a close supervision of public attitudes towards the dissidents, which could range from disapproval to sympathy, and ultimately to willingness to participate in the resistance. Under these conditions, full access (and thus surveillance), paired with a highly selective coercive campaign directed against the core dissidents is likely to be most successful.

Shutting down or limiting all access to information and communication would prove ineffective for multiple reasons. First and foremost, it would weaken the government's ability to target the dissidents responsible for generating a threat. However, there are other far-reaching repercussions: restrictions or bans on (social) media generate substantial negative press that is likely to draw more attention to the cause of the dissidents, thus making it harder for the government to quietly clamp down on the unrest (see, for example Tufekci, 2014). Lastly, public support from those originally in favor of the status quo is likely to suffer where basic access to network services is curtailed. In consequence, where governments plan on conducting repressive campaigns that are targeted against

certain individuals, groups, or organizations, the free exchange of information is likely to be more advantageous than the limitation thereof.

- **Empirical Expectation 1:** Government provisions for the free exchange of information are likely to be positively correlated with a targeted coercive response tactic.

8.3.2 Incentives for censorship

In situations where the source of the threat is unclear and/or a substantial proportion of the population is deemed threatening to the government's political authority, the trade-off between surveillance and censorship produces a different solution. Assuming that, for example, the majority of the population is known to be in favour of an alternative ruling power and is willing to mobilize against the status quo, the repressive response required by desperate government is likely to aim at signalling strength and resolve (Downes, 2007). Put bluntly, where everyone is seen as a potential threat, no nuances in the surveillance of individuals or groups is required, as the whole population has been identified as a potential target, regardless of their online activities. On the contrary, a free exchange of information would mean providing a free infrastructure for the opposition to organize their rebellion effectively via the internet.

In situations where the most effective strategy of repression is anticipated to be an extreme scorched-earth-policy (see Downes, 2007), governments will reap the highest benefit from restricting access to communication and information. In situations where the government cannot afford to implement such an extreme policy, but nevertheless perceives its position of power to be so severely threatened that forceful, untargeted retaliation is attempted, restricting information access offers yet another form of indiscriminate punishment.

- **Empirical Expectation 2:** Government restrictions on the free exchange of information are likely to be positively correlated with a less targeted (more untargeted) coercive response tactic.

8.4 Data and empirical strategy

8.4.1 Regional network accessibility in Syria

To measure regional network accessibility in Syria, I make use of survey data collected by the Syria Digital Security Monitor (SDSM), a project funded by the Ottawa-based SecDev Foundation. Since June 2013, SDSM has surveyed all Syrian districts on a biweekly[1] basis in order to establish the degree of digital accessibility across the country. The survey asks respondents to separately rate their ability to use the internet (distinguishing between ADSL, 2G, 3G), landlines and mobile phones on a four-point scale, where 1= general availability, 2 = available often, 3= intermittent availability, and 4 = no availability. To ensure comparability, SDSM attempts to survey the same set of respondents in every wave, but also makes use of social media sources.[2] To obtain a standardised unit of analysis, the accessibility measures are aggregated to the governorate level,

[1] For some months, only one survey is available, not two.
[2] Personal communication with SecDev Foundation staff.

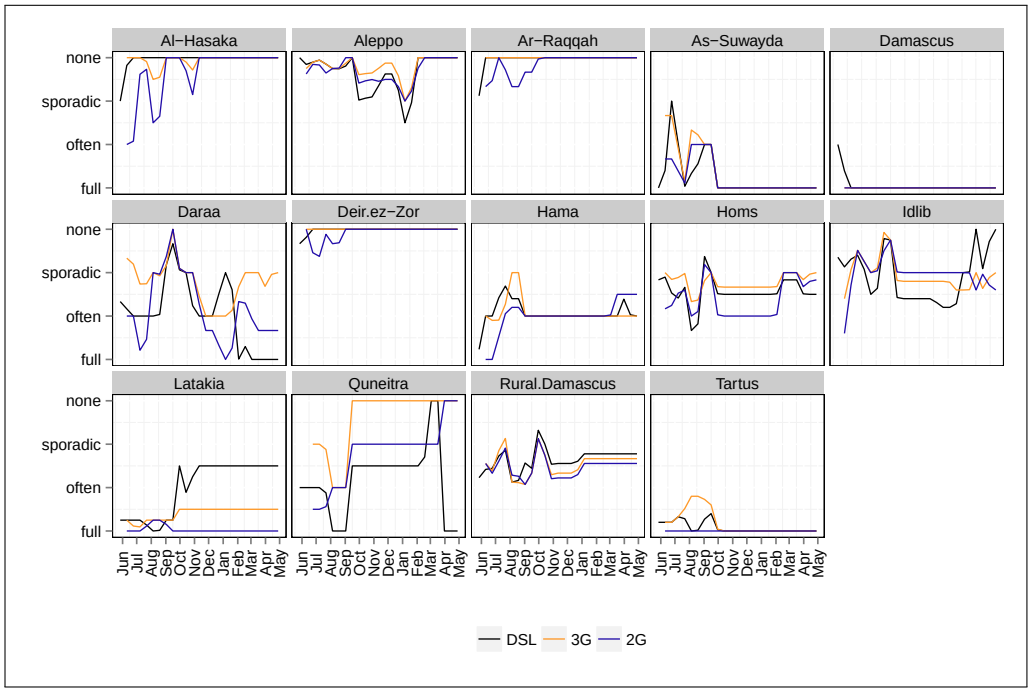


Figure 8.1. Internet (DSL, 3G, and 2G) accessibility by Syrian governorate, June 2013 - April 2014.

measured in biweekly intervals. I test the effect of connectivity using the three measures for internet connectivity, which are access to DSL-internet, as well as mobile access to 2G and 3G networks. As the availability of mobile phone networks has been the subject of recent research, I run additional test with this measure (Shapiro and Weidmann, forthcoming; Pierskalla and Hollenbach, 2013), but do not expect a significant effect here. The theoretical expectations formulated above hold no prediction of the effect of landlines, as these long predate the internet era. I therefore expect no significant effect between the availability of landline telephone access and repressive strategies.

Figure 8.1 plots the level of internet accessibility (DSL, 3G and 2G) by governorate for the time period of this study, June 2013 - April 2014. Where the lines spike, no or only little internet access is available. It is important to note that – with exception of Al-Hasaka and Ar-Raqqah – all regions that witness temporary inaccessibility throughout this time period, regain connectivity. If reductions in internet access were tied to technical failures stemming from irreparable damage by destruction, the loss of access would be irreversible and no return to access would be discernible. Instead, we see that in the most contentious regions (among others: Aleppo, Rural Damascus, Homs) connectivity varies considerably over time.

8.4.2 Classifying targeted and untargeted repression

Quantifying variations in a government's repressive strategy is challenging, in particular when the objects of interest are patterns that move beyond scale. The operational definition for targeted and untargeted violence used here builds on

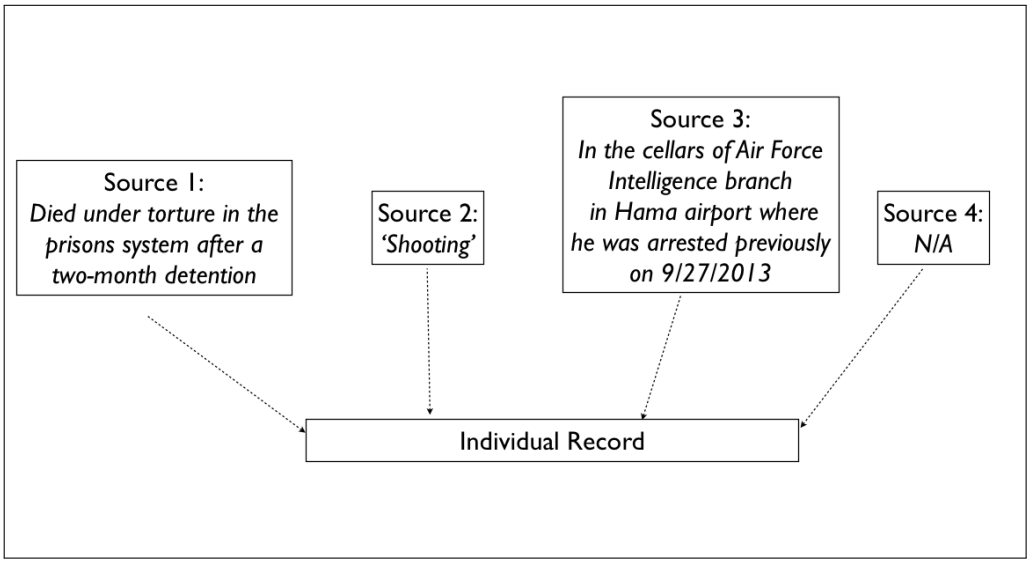


Figure 8.2. Assembling information on record details.

work by Kalyvas (2006), Steele (2009), and Wood (2010). In this context, state violence is defined as targeted if the victim was killed either due to individual (dissident, defector, critical journalist) or collective (party, ethnicity) characteristics. All incidences where the victim was not selected based on individual or collective characteristics are assumed to be untargeted violence. Since it is impossible to measure intent, I rely on descriptive information regarding the circumstances of violence to infer the probable intent. I use supervised machine-learning to classify over 60,000 aggregated reports on individual killings that were committed by the Syrian Regime (and pro-government forces) between June 2013 and April 2014 (see Price, Gohdes and Ball, 2014).

To arrive at an accurate number of reported killings for the period concerned, I combine information on lethal violence in Syria collected by different human rights groups, as presented in Chapter 5. The groups' records were pooled to one large dataset. Each record in this 'pool' was then compared to every other record in order to identify records that match across name, date of death and location (governorate)(Price, Gohdes and Ball, 2014). De-duplicating and matching different data sources on violence offers two important analytical advantages: First, it reduces the probability of over-representing violent events covered by multiple groups, a known problem of event-data sources, as discussed in Chapter 5. Second, it allows for the triangulation of all auxiliary information that comes with each of these records. Figure 8.2 presents a model of this triangulation process for an individual record: Assuming that a violent event was documented by three of the four groups, we can combine the notes on death circumstances from these three sources to arrive at a more nuanced understanding of how (and possibly why) this person was killed by the government. These combined, or 'aggregated' texts form the basis on which the classification of records is conducted.

Based on a training set of 2,000 randomly selected records I classify by hand,[3] the model is trained to classify each record as either a targeted killing or an untargeted killing. A third category was classified which includes all victims killed by non-government forces. These are predominantly attributed to ISIS (Islamic State of Iraq and the Levant) forces, and a few were also attributed to PKK/PYK (Kurdish) forces. All records in this final category were dropped from the analysis.

In the hand-coded training set, records are classified as **targeted** killings if the circumstances described in the aggregated report 1) indicate that the victim was selected based on his/her specific characteristics *(e.g. 'killed because he refused to [...]', 'targeted while protesting [...]', 'dissent')* and/or 2) indicate that the method of killing was of a selective nature *(e.g. executed by sniper, hanging, beheading, set afire)*, and/or 3) the method of killing was accompanied by other violations of a selective nature *(e.g. arrest, detention, prison, 'found with hands/legs tied')*. Records are classified as **untargeted** killings if the circumstances described in the aggregated report 1) indicate that the victim was not selected based on his/her specific characteristics *(e.g. 'stepped on a landmine')*, and/or indicate that the method of killing was not selective *(e.g. explosion, bombing, shelling, mortar, chemical, toxic cases)*, and/or 3) the method of killing was not accompanied by other targeted violations.

The results presented here use the support vector machine (SVM) learning algorithm (Cristianini and Shawe-Taylor, 2000; Feinerer, Hornik and Meyer, 2008).[4] Figure 8.3 presents a summary of the classified records. Of the 60,904 records, 825 could not be classified because no additional information was available from any of the sources. 1780 of the recorded killing were perpetrated by either the PKK or ISIS. The unidentifiable records and those perpetrated by PKK and ISIS were omitted from the analysis. The majority of the records collected by the four human rights groups indicate untargeted violence, and more than 8000 are classified as targeted instances of state repression. For each location and time point, I establish the number of targeted killings $targ_{it}$, and the number of untargeted killings $untarg_{it}$, which together form the overall number of regime fatalities n_{it}.

8.4.3 Estimating reliable levels of targeted and untargeted killings

Variations in reporting can be a serious problem when using 'raw' event data (Weidmann, 2014*b*), in particular when attempting to compare patterns of categories of violence that occurred under different circumstances. As such there are two important reasons why this study cannot make use of simple event counts. First, the likelihood that targeted and untargeted violent events are reported with a different probability is high. Second, the main variable of interest - variations in information accessibility - is likely to heavily influence the ability to report, which by consequence would lead to reports of violence being endogenous to the level of information accessibility.

[3]The information provided by the four human rights groups was translated from Arabic into English using *Goslate*, a Free Google Translate API for Python.

[4]The classification process was performed using the RTextTools package written by Jurka et al. (2012). A variety of different algorithms were tested, including Random Forest, maximum entropy and glmnet, and the results were all comparable. The SVM algorithm consistently provided the highest overall accuracy, which is why I only use the SVM classifications in this chapter.

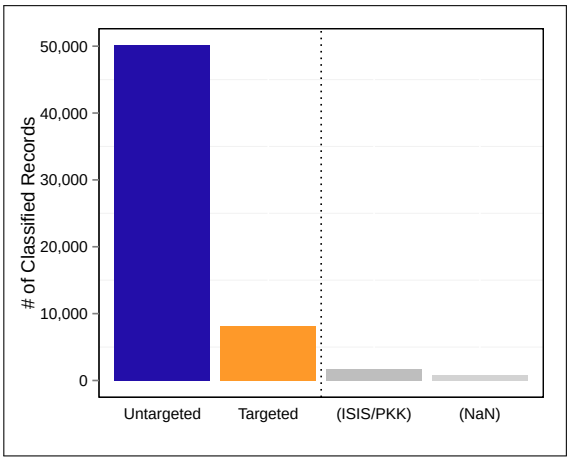


Figure 8.3. Summary of classified records.

Using multiple recapture estimation, as introduced in Chapter 6, I predict the number of unreported targeted and unreported untargeted killings for every governorate, for every time period covered by an SDSM survey (see also Baillargeon and Rivest, 2007; Gohdes, 2014). Instead of using the counts in each category generated from the classification, I use the estimated count for both targeted and untargeted violence, for each observation of accessibility in each Syrian governorate between June 2013 and April 2014. Since the survey was conducted on a bi-weekly basis and there are 14 governorates, I arrive at an N=350.

To account for uncertainty in the estimation process, I repeat the empirical analysis with the upper and the lower bound of each estimated count (also know as the 95% confidence intervals). The analyses are reported in the Appendix in Figure 5 and Figure 6. The results are comparable.

8.5 Results

The analysis defines a government's coercive tactic to consist of two components, which are the number of targeted and the number of untargeted killings. To account for both scale and proportion between the two, I fit a beta-binomial generalised linear model which is generally used for clustered data with the form $\{n, m\}$, where n is the size of the entire cluster, and m is defined as the number of successes in n, so $m \leq n$ (see Lesnoff and Lancelot, 2013: 3-7). In this study, n_{it} is the number of overall fatalities committed by the regime at a given time and place (targeted + untargeted), and m_{ii} is the number of those fatalities that were targeted killings. The dependent variable is defined as $y = \frac{m}{n}$. For a given observation (n, m), the model is

$$m|\lambda, n \sim Binomial(n, \lambda), \tag{8.1}$$

where λ follows a Beta distribution $Beta(a1, a2)$. I opt against a conventional fixed-effects approach to account for unobserved heterogeneity that might result from the panel structure of the data, since the descriptive statistics suggest that the between-variation is likely to drive a substantial amount of the effect of

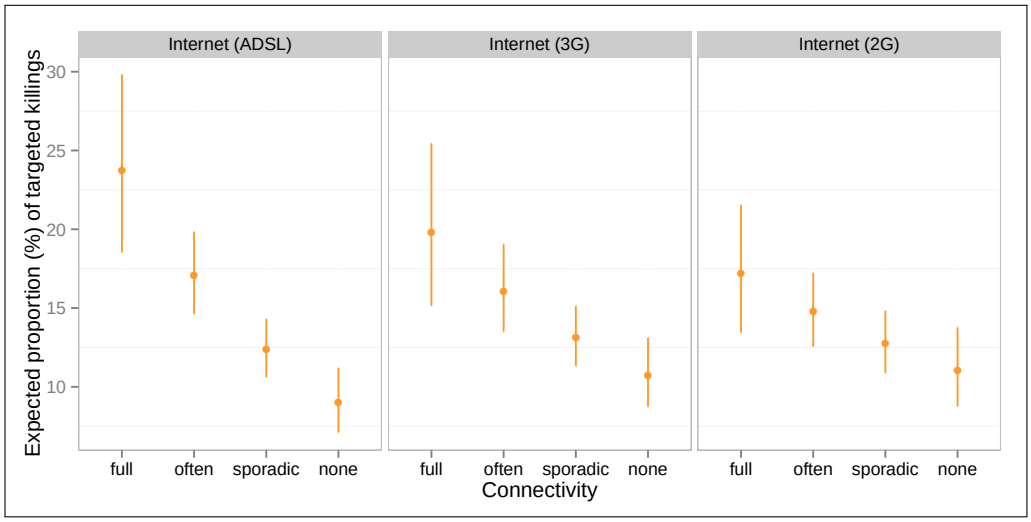


Figure 8.4. Expected proportion of targeted killings, given internet accessibility.

interest. To account for heterogeneity, I estimate separate dispersion parameters for each governorate. Initial comparisons of AIC/BIC model fit criteria indicate that the heterogeneous dispersion models provide the best model fit.

Figure 8.4 presents the simulated expected proportion (with 95% confidence intervals) of targeted killings (in %) for each degree of connectivity, by type of connection. The left panel presents the effect of access to regular internet, where the expected proportion of selectively targeted killings by the government is more then 20% (with a confidence interval of 16.5% and 26.9%). According to this model, one in five victims are purposefully selected to be killed by the government where access to network services is allowed. As the level of connectivity decreases, this proportion drops significantly. Where access to the internet is fully censored, the model predicts that only about 8% (with a confidence interval of 7.2% to 10.6%) of all victims have been targeted based on individual or collective characteristics, which means that 9 out of 10 killed have been indiscriminately attacked. This large and significant difference in coercive strategies offers empirical support for the expectations formulated in this study: The strategy of government coercion varies significantly with the level of internet connectivity provided. The other two measures of internet accessibility offer similar results. Anecdotal evidence from Syrian citizens suggests that the restrictions placed on their telecommunications system are commonly understood to be a part of 'mass-punishment' by the government, which is becoming increasingly desperate to regain control over 'lost' parts of the country.[5] Figure 8.5 presents the effect of mobile phone and landline accessibility on varying strategies of state repression. Mobile phone connectivity seems to follow the same pattern of internet connectivity, but, as expected, the accessibility of landlines bears no significant relationship with the government's coercive strategy. To facilitate the interpretation of the results, Figure 8.6 presents the simulated expected change (in percentage points) in the proportion of targeted killings, when moving from no access to full connectivity. Where the 95% confidence interval does not include zero, the expected change is significant. The

[5]Personal interview with two Syrian activists, and staff of the Syrian Digital Security Monitor.

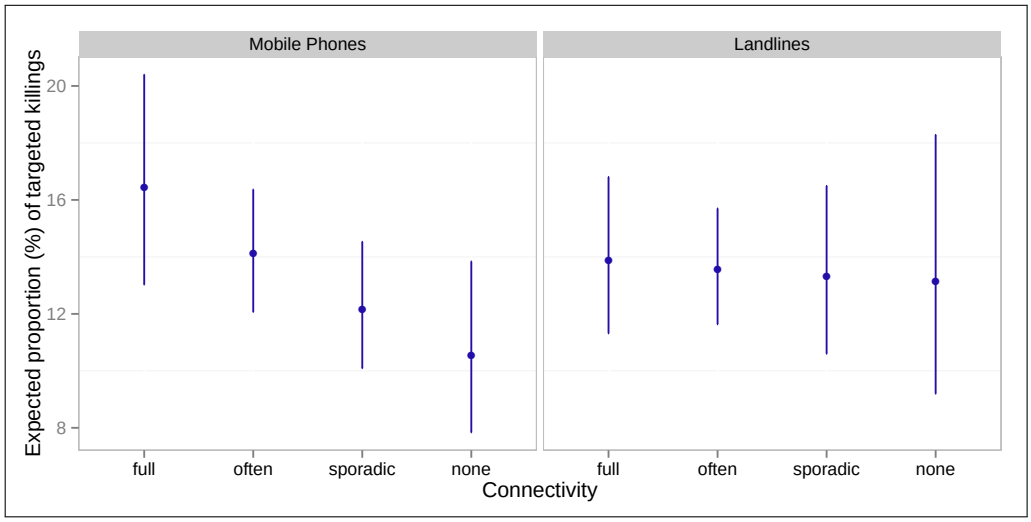


Figure 8.5. Expected proportion of targeted killings, given internet accessibility.

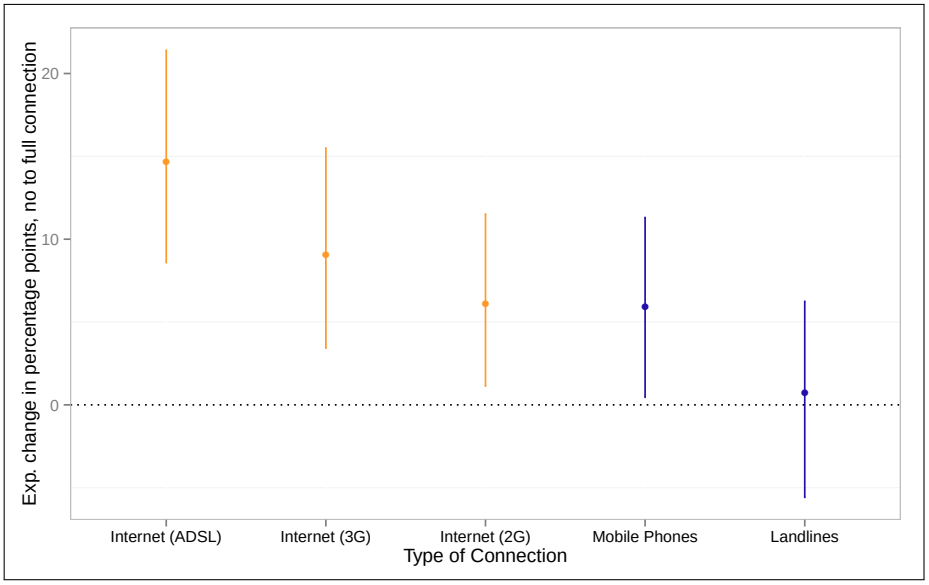


Figure 8.6. First difference: change in percentage of targeted killings (no access to full connection).

first-difference graph improves the illustration of the findings: as the availability of internet access moves from none to full, the percentage of victims that were purposefully targeted increases significantly for all types of connections, except (as expected) for landlines. Where connections are restricted, the government follows a coercive strategy that is significantly more indiscriminate.

8.6 Summary

This chapter provided the most rigorous test for the overall theoretical argument of the dissertation: governments spatially and temporally vary their strategies

of internet control and adapt their violent repressive strategy accordingly. The analysis provides solutions for a number of challenges to measuring both state control of the internet and variations in the strategies of state repression. Survey data on the level of accessibility of the internet was used to gauge the state's implementation of network controls. The Syrian government has a demonstrated history of using different techniques of digital surveillance to spy on its own population, which is why it is plausible to infer that where the internet is accessible, it is using it to monitor the opposition. Limits to accessibility in turn indicate strategic restrictions, not least because the level of connectivity has strongly fluctuated over time and regions, indicating that lack of access is likely to be due to technical failures.

To measure repressive strategies, I used supervised machine-learning algorithms to classify over 60,000 records of state killings according to their event circumstances. Capture-recapture models were then used to estimate the total number of killings in both categories for each governorate and time period in the dataset on network restrictions. The analysis of the effect of varying levels of network control on the proportion of targeted (versus untargeted) state killings reveals that higher levels of connectivity are significantly related to an increase in the proportion of targeted repression. In contrast, little or no access to the internet is significantly linked to indiscriminate campaigns of violence. These results offer systematic support for the central theoretical arguments presented in this study. A discussion of the implications of these findings, and the findings presented in previous chapters, is delivered in the final chapter of this dissertation.

9 Conclusion

The digital revolution presents governments who fear for their political survival with a dilemma. On the one hand, the exponential increase in 'private', user-generated online content offers seemingly endless possibilites for surveillance, making the spying, tracking, profiling, and ultimately the isolating and targeting of individuals deemed threatening to the regime cheap and efficient. On the other hand, civilian uprisings from Damascus to Cairo have shown that social media has the potential for collectively mobilizing protesters and spreading alternative views on political actions in a way that was previously neither possible nor foreseeable. When the citizens of Hong Kong took to the streets in September 2014, smartphones were the central medium by which their movement was organised, promoting a common spirit, coordinating events, and ensuring the availability of basic necessities, such as food, water and sanitation (Shirky, 2014). Digital images of hundreds of thousands of protesters holding up illuminate smartphone screens instead of candles travelled around the world. The Chinese government responded by censoring any references to the protests on its social media platforms to prevent the spread of information to the mainland, but has so far refrained from actually cutting internet accessibility (Olesen, 2014). Pro-democracy websites were found to have malware embedded that compromised the computers and phones of those visting them, thereby facilitating full digital surveillance (Adair, 2014). Along with increased digital empowerment of ordinary citizens, states have also increased their own abilities to tamper with, and temporarily limit, centrally-provided network services, by making use of increasingly sophisticated technology.

The evidence presented in this dissertation clearly shows that internet control is pervasive and unmistakably linked to states' larger repressive strategies. The global analysis of network disruptions and states' abuse of physical integrity rights presented in Chapter 4 shows that even when taking into account the main predictors of state repression, disruptive internet controls are significantly and positively associated with increased state terror. The case examples demonstrated the variability in network controls and the subtle ways in which both internet censorship and surveillance are used to inform states' coercive strategies.

Indeed, even in countries like Ethiopia, where only a small fraction of citizens make use of the internet, governments have installed pervasive controls of the digital sphere. The quantitative case studies in Chapters 7 and 8 provided extensive and disaggregated evidence for the instrumental value governments ascribe to controlling internet accessibility in times of civil conflict. The results discussed in Chapter 7 suggest that the Syrian government implemented full-blown blackouts of the internet in conjunction with larger military offensives against the opposition and civilians supporting the opposition. Through the use of a number of innovative statistical methods, I was able to eliminate important competing explanations for this empirical association, such as the plausible alternative explanation that shutdowns were intended to hide atrocities from the outside world. Chapter 8 revealed that the nature of state violence is linked to variations in network accessibility, when comparing the level of targeted and untargeted lethal violence used by the Syrian regime in different regions of the country.

In the concluding section, I will now briefly summarise the main theoretical and empirical implications that this dissertation provides, and discuss the broader ramifications for policy makers and society at large.

9.1 Theoretical contribution

9.1.1 Surveillance, censorship, and violence after the digital revolution

This is the first study to theoretically link the use of purposely-implemented internet controls to the type and scale of state-sanctioned violence repression. The availability of intelligence, and the manipulation of information access have both featured prominently in theories of state repression and authoritarian rule (Kalyvas, 2006; Kern and Hainmueller, 2009; Davenport, 2010; Kim, Whitten-Woodring and James, 2014). The rapid shift in the nature, access to, and producers of information in the digital age call for theoretical approaches that account for these changes. This study adds to our understanding of how online information access and the power of digital intelligence gathering via surveillance, inform states' opportunities and willingness to respond to domestic unrest with violent repression. Future data collection efforts on states' use of digital surveillance tools will be required to test these mechanisms at a global level, and will provide important insights into specific types and applications of surveillance in different political contexts.

9.1.2 Network control and military warfare

Evidence presented in Chapter 7 demonstrates the potential of network restrictions to constitute a tactic within larger military offensives. The analyses of internet shutdowns and state killings suggest that regimes use large-scale disruptions selectively and purposely in conjunction with concerted repressive offensives against the opposition. Not all military offensives are accompanied by outages, but when access is denied, there is a significant rise in the number of lives lost.

Syria represents the first conflict that has been meticulously followed and fuelled by a vast online audience: by the opposition fighters and supporters, by regime forces and their supporters, and by the outside world at large. The increasing importance of establishing control over online content and access to the internet, is likely to exert a growing appeal for regimes eager to adjust their repertoire of repressive tools in dealing with new digital threats to the status quo.

9.2 Methodological contribution

9.2.1 Integration and classification of high-quality data on government coercion

The new database on state killings in Syria presented in Chapter 5 addresses the problem of over-counting in the collection of event data on political violence. Linking records provided by five different sources on the occurrence of lethal violations committed by the regime, ensures that every incident is only counted once. The fact that each record can be traced back to the original sources in which it was found provides a number of opportunities for researchers to gain a deeper understanding of how the data were collected, and how this process varies over temporal and spatial units. Linking information from multiple sources also allowed me to combine detailed evidence of all sources on circumstances under which the killing was perpetrated. The circumstantial evidence was used to classify victims with the help of supervised machine-learning algorithms, according to the nature of their death. The classification into targeted and untargeted state killings provides an improved methods of measuring strategies of repression, and demonstrates the great potential machine-learning approaches offer for the empirical analysis of political violence.

9.2.2 Addressing bias in documented event-data

This dissertation could not have answered the question of how internet control is linked to repression without accounting for potential bias in observable measures of state violence. In general, research on the relationship between varying forms of information control and conflict is likely to suffer from endogenous measures of information access and violence, as our knowledge of violence depends on the availability of information. The statistical solution presented in this dissertation has already been used in a variety of other scientific fields, ranging from demography to epidemiology, but has to date found few applications in the social sciences. The analyses presented in this dissertation are among the first to use corrected statistical predictions of violence for the empirical analysis of contentious dynamics.

9.3 Policy implications

9.3.1 Government accountability

Evidence presented here shows clearly that governments use internet control to inform their use of violence against their own citizens. This has impor-

tant ramifications for policies aimed at protecting human rights and ensuring accountability for those who abuse them.

Foreign governments and the international community should begin to understand state-led disruptions of internet accessibility as a serious signal; they should strongly and swiftly condemn any such occurrence. Disruptions should simply be viewed as a means of stifling the opposition's ability to communicate. They should be understood as a clear signal of repressive intent by a government set on maintaining its political power at all costs.

On a more hopeful note, the rise in alternative means of obtaining internet accessibility, for example by using satellite connections or connections from neighbouring countries, means that it has become increasingly difficult for governments to impose true isolation on a whole population. A recent study by Dyn Research (formerly Renesys) demonstrates that countries with a centralised telecommunications sector are at a much higher risk of being subject to internet disruptions, as the central structure requires far fewer steps in order to disconnect all digital entry-points to the domestic cyberspace (Cowie, 2014). The report finds that countries with fewer than 10 internet service providers are particularly vulnerable, including Myanmar, Syria, Yemen, and Rwanda. Countries with more than 40 different service providers, such as the United States, Germany, or France, are highly unlikely to suffer a full blackout (ibid.). Policy makers should therefore actively encourage the diversification of the telecommunications market in countries with a known history of state terror.

With the increase of surveillance software being used by governments to spy on their citizens, the opportunities for identifying dissidents and willingness to eliminate them has risen dramatically. As discussed in detail in Chapter 3, the majority of surveillance software used by despotic governments is exported from companies located in the European Union and the United States (see e.g. Raoof, 2011; Wagner and Guarnieri, 2014). State abuse of digital surveillance software should give policy makers in democratic countries serious reason to consider the careful regulation of exports on these types of software. Analogous to policies restricting the export of arms to governments known to turn these weapons against their own population, policy reform is needed to address the growing abuse of digital spyware and malware. Initiatives such as the *Coalition Against Unlawful Surveillance Exports*[1] and the Global Surveillance Monitor launched by *Privacy International*[2], as well as the *Open Rights Group*[3], are already actively pursuing this topic.

9.3.2 Security implications for citizens and activists

If we move away from the perspective of repressive governments, the findings of this dissertation offer important implications for citizens and activists caught in the process of challenging the political status quo. Reliance on the internet to collectively organize sustainable campaigns and resistance movements offers many advantages, but it is clearly a double-edged sword: digital fingerprint – from emails to call histories and Facebook *likes* – mean that activists are liable to be placed under close surveillance, long before their campaigns become

[1] `http://www.globalcause.net/`

[2] `https://www.privacyinternational.org/campaigns/global-surveillance-monitor`

[3] `https://www.openrightsgroup.org/`

publicly known. The acquisition of sophisticated knowledge of ways to securely communicate, work, live, and travel without leaving a clear 'paper-trail' is no longer simply recommendable to a select number of clandestine dissidents. It may well become a matter of survival for all members of an opposition, because the findings of this study suggest that where a government does allow the free exchange of information, it is more likely to be engaged in selective targeting and killing of its citizens. The defence of every individual's basic rights to privacy and confidential communication has never been more essential than in the current digital age.

Bibliography

Adair, Steven. 2014. "Democracy in Hong Kong Under Attack." *Volexcity Blog* 9 October.

Aday, Sean, Henry Farrell and Marc Lynch. 2010. "Blogs and Bullets: New media in contentious politics." *United States Institute of Peace Peaceworks* 65.

Amnesty International. 2012. "Syria: Shutting down of internet and mobile networks alarming development." *Amnesty International Press Release* 29 November.

Arsenault, Amelia, Sheldon Himelfarb and Susan Abbott. 2011. "Evaluating Media Interventions in Conflict Countries." *United States Institute of Peace* .

Baillargeon, Sophie and Louis-Paul Rivest. 2007. "Rcapture: Loglinear Models for Capture-Recapture in R." *Journal of Statistical Software* 19(5):1–31.

Balcells, Laia. 2010. "Rivalry and Revenge: Violence against Civilians in Conventional Civil Wars." *International Studies Quarterly* 54(2):291–313.

Ball, Patrick, Jana Asher, David Sulmont and Daniel Manrique. 2003. "How Many Peruvians Have Died? An Estimate of the Total Number of Victims Killed or Disappeared in the Armed Internal Conflict Between 1980 and 2000." Washington, DC: Report to the Peruvian Commission for Truth and Justice (CVR).

Ball, Patrick, Paul Kobrak and Herbert F. Spirer. 1999. *State Violence In Guatemala, 1960-1996: A Quantitave Reflection*. American Association for the Advancement of Science.

Ball, Patrick, Wendy Betts, Fritz Scheuren, Jana Dudukovich and Jana Asher. 2002. "Killings and Refugee Flow in Kosovo March - June 1999." Washington, DC: A Report to the International Criminal Tribunal for the Former Yugoslavia.

Bamford, James. 2014. "Edward Snowden: The Untold Story." *WIRED Magazine* August.

Baum, Matthew A. and Yuri M. Zhukov. forthcoming. "Filtering Revolution: Reporting Bias in International Newspaper Coverage of the Libyan Civill War." *Journal of Peace Research* XX.

Beaumont, Peter. 2013. "Syrian missiles similar to 'ones used in previous chemical weapons attacks'." *The Guardian* 22 August.

Beiser, Elana. 2013. "Syria, Iraq, Egypt most deadly nations for journalists." *Committee to Protect Journalist Special Report* 30 December.

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

Bennett, W. Lance and Alexandra Segerberg. 2013. *The logic of connective action: Digital media and the personalization of contentious politics*. New York, NY: Cambridge University Press.

Bhavnani, Ravi, Dan Miodownik and Hyun Jin Choi. 2011. "Three Two Tango: Territorial Control and Selective Violence in Israel, the West Bank, and Gaza." *Journal of Conflict Resolution* 55(1):133–158.

Bishop, Yvonne M. M., Stephen Fienberg and Paul H. Holland. 1975. *Discrete Multivariate Analysis: Theory and Practice*. Cambridge, MA: MIT Press.

Blattman, Christopher and Edward Miguel. 2010. "Civil war." *Journal of Economic Literature* 48(1):3–57.

Brandt, Patrick T., John T. Williams, Benjamin O. Fordham and Brain Pollins. 2000. "Dynamic Modeling for Persistent Event-Count Time Series." *American Journal of Political Science* 44(4):823–843.

Breuer, Anita, Todd Landman and Dorothea Farquhar. forthcoming. "Social media and protest mobilization: evidence from the Tunisian revolution." *Democratization* XX(0):1–29.

Browning, John G. 2013. "Democracy unplugged: social media, regime change, and governmental response in the Arab Spring." *Michigan State International Law Review* 21:1.

Brownstone, Andy. 2011. "Meet the Libyan rebels on the front line." *BBC Newsbeat* 24 May.

Brück, Tilman, Patricia Justino, Philip Verwimp and Alexandra Avdeenko. 2010. Identifying Conflict and Violence in Micro-Level Surveys. HiCN Working Paper 79 Households in Conflict Network (HiCN).

Bueno de Mesquita, Bruce, Alastair Smith, Randolph M. Siverson and James D. Morrow. 2003. *The logic of political survival*. Cambridge, MA: MIT Press.

Carey, Sabine C. 2009. *Protest, repression and political regimes: an empirical analysis of Latin America and sub-Saharan Africa*. Routledge.

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

Carey, Sabine C. and Mark Gibney. 2010. *The Politics of Human Rights*. Cambridge Univeristy Press.

Castells, Manuel. 2012. *Networks of outrage and hope : social movements in the internet age*. 1 ed. Cambridge, MA: Polity Press.

Chojnacki, Sven, Christian Ickler, Michael Spies and John Wiesel. 2012. "Event data on armed conflict and security: New perspectives, old challenges, and some solutions." *International Interactions* 38(4):382–401.

Chowdhury, Mridul. 2008. "The role of the Internet in Burma's Saffron Revolution." *Berkman Center Research Publication* (2008-8).

Cingranelli, David L. and David L. Richards. 1999. "Measuring the level, pattern, and sequence of government respect for physical integrity rights." *International Studies Quarterly* 43(2):407–417.

Cingranelli, David L. and David L. Richards. 2010. "The Cingranelli and Richards (CIRI) human rights data project." *Human Rights Quarterly* 32(2):401–424.

Cohen, Noam. 2009. "Twitter on the barricades: Six lessons learned." *The New York Times* 20 June.

Committee to Protect Journalists. 2014. *Attacks on the Press: Journalism on the World's Front Lines, 2014 Edition*. John Wiley & Sons.

Condra, Luke N. and Jacob N. Shapiro. 2012. "Who Takes the Blame? The Strategic Effects of Collateral Damage." *American Journal of Political Science* 56(1):167–187.

Conrad, Courtenay R. 2011. "Constrained Concessions: Beneficent Dictatorial Responses to the Domestic Political Opposition." *International Studies Quarterly* 55(4):1167–1187.

Conrad, Courtenay R. and Will H. Moore. 2010. "What Stops the Torture?" *American Journal of Political Science* 54(2):459–476.

Conrad, Courtenay R. and Will H. Moore. 2012. "Political Institutions, Plausible Deniability, and the Use of Stealth Torture.".

Cormack, Richard M. 1989. "Log-Linear Models for Capture-Recapture." *Biometrics* 45(2):pp. 395–413.

Cowie, Jim. 2014. "Syria, Venezuela, Ukraine: Internet Under Fire." *Renesys Blog* 26 February.

Cristianini, Nello and John Shawe-Taylor. 2000. *An introduction to support vector machines and other kernel-based learning methods*. Cambridge university press.

Dafoe, Allan and Guadalupe Tunón. 2014. "Placebo Tests for Causal Inference." ISA Annual Meeting Paper.

Danneman, Nathan and Emily Hencken Ritter. 2014. "Contagious Rebellion and Preemptive Repression." *Journal of Conflict Resolution* 58(2):254–279.

Davenport, Christian. 1995. "Multi-Dimensional Threat Perception and State Repression: An Inquiry into Why States Apply Negative Sanctions." *American Journal of Political Science* 39(3):683–713.

Davenport, Christian. 1999. "Human Rights and the Democratic Proposition." *Journal of Conflict Resolution* 43(1):92–116.

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

Davenport, Christian. 2007*b*. "State Repression and the tyrannical Peace." *Journal of Peace Research* 44(4):485–504.

Davenport, Christian. 2010. *Media bias, perspective, and state repression: the black panther party*. New York: Cambridge University Press.

Davenport, Christian. forthcoming. *How Social Movements Die. Repression and Demobilization of the Republic of New Africa*. New York: Cambridge University Press.

Davenport, Christian and II Armstrong, D. A. 2004. "Democracy and the Violation of Human Rights: A Statistical Analysis from 1976 to 1996." *American Journal of Political Science* 48(3):538–554.

Davis, David R. and Michael D. Ward. 1990. "They Dance Alone: Deaths and the Disappeared in Contemporary Chile." *Journal of Conflict Resolution* 34(3):449–475.

De Mesquita, Bruce Bueno, Feryal Marie Cherif, George W Downs and Alastair Smith. 2005. "Thinking inside the box: A closer look at democracy and human rights." *International Studies Quarterly* 49(3):439–458.

Deibert, Ronald J. 2008. *Access denied: The practice and policy of global internet filtering*. Cambridge, MA: MIT Press.

Deibert, Ronald J. 2009. The geopolitics of internet control: Censorship, sovereignty, and cyberspace. In *The Routledge handbook of internet politics*, ed. Andrew Chadwick and Philip N. Howard. Oxon: Routledge pp. 323–336.

Deibert, Ronald J. 2010. *Access controlled: The shaping of power, rights, and rule in cyberspace*. The MIT Press.

Deibert, Ronald J and Rafal Rohozinski. 2010. Control and subversion in Russian cyberspace. In *Access controlled: The shaping of power, rights, and rule cyberspace*, ed. Ronald Deibert. Cambridge, MA: MIT Press pp. 15–34.

Dewey, Taylor, Miriam Marks Juliane Kaden, Shun Matsushima and Beijing Zhu. 2012. "The Impact of Social Media on Social Unrest in the Arab Spring." *Final Report prepared for: Defense Intelligence Agency* .

Diamond, Larry. 2010. "Liberation technology." *Journal of Democracy* 21(3):69–83.

Diamond, Larry and Marc F. Plattner. 2012. *Liberation technology: Social media and the struggle for democracy*. Johns Hopkins University Press.

DiNucci, Darcy. 1999. "Design & New Media: Fragmented Future-Web development faces a process of mitosis, mutation, and natural selection." *Print magazine* 53:32–35.

Downes, Alexander. 2007. "Draining the sea by filling the graves: investigating the effectiveness of indiscriminate violence as a counterinsurgency strategy." *Civil Wars* 9(4):420–444.

Downes, Alexander B. 2008. *Targeting civilians in war*. Ithaca, NY: Cornell University Press.

Driessen, Benedikt, Ralf Hund, Carsten Willems, Christof Paar and Thorsten Holz. 2012. "Don't Trust Satellite Phones: A Security Analysis of Two Satphone Standards." *IEEE Symposium on Security and Privacy* :128–142.

Dube, Arindrajit, Ethan Kaplan and Suresh Naidu. 2011. "Coups, Corporations, and Classified Information." *Quarterly Journal of Economics* 126(3):1375–1409.

Earl, Jennifer. 2011. "Political Repression: Iron Fists, Velvet Gloves, and Diffuse Control." *Annual Review of Sociology* 37(1):261–284.

Earl, Jennifer, Andrew Martin, John D. McCarthy and Sarah A. Soule. 2004. "The Use of Newspaper Data in the Study of Collective Action." *Annual Review of Sociology* 30:65–80.

Earl, Jennifer and Katrina Kimport. 2011. *Digitally enabled social change: Activism in the internet age*. Cambridge, MA: MIT Press.

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

Eck, Kristine and Lisa Hultman. 2007. "One-Sided Violence Against Civilians in War." *Journal of Peace Research* 44(2):233–246.

Edmond, Chris. 2011. "Information manipulation, coordination, and regime change." *NBER Working Paper* 17395.

Else, Liz. 2012. "The revolution will be tweeted." *The New Scientist* 6 February.

Escribà-Folch, Abel. 2013. "Repression, political threats, and survival under autocracy." *International Political Science Review* 34(5):543–560.

Fariss, Christopher J. 2014. "Respect for Human Rights has Improved Over Time: Modeling the Changing Standard of Accountability." *American Political Science Review* 108(2):297–318.

Fariss, Christopher J. and Keith E. Schnakenberg. 2014. "Measuring Mutual Dependence between State Repressive Actions." *Journal of Conflict Resolution* 58(6):1003–1032.

Fein, Helen. 1995. "More Murder in the Middle: Life-Integrity Violations and Democracy in the World, 1987." *Human Rights Quarterly* 17(1):170–191.

Feinerer, Ingo, Kurt Hornik and David Meyer. 2008. "Text Mining Infrastructure in R." *Journal of Statistical Software* 25(5):1–54.

Fienberg, Stephen. 1972. "The Multiple recapture census for closed populations and incomplete 2^k contingency tables." *Biometrika* 59(3):591–603.

Francisco, Ronald A. 1995. "The Relationship between Coercion and Protest: An Empirical Evaluation in Three Coercive States." *Journal of Conflict Resolution* 39(2):263–282.

Frantz, Erica and Andrea Kendall-Taylor. 2014. "A dictator's toolkit: Understanding how co-optation affects repression in autocracies." *Journal of Peace Research* 51(3):332–346.

Gallagher, Sean. 2012. "Updated: Paint it black—How Syria methodically erased itself from the 'Net." *Ars Technica* 1 December.

Galperin, Eva and Morgan Marquis-Boire. 2012. "Fake YouTube Site Targets Syrian Activists With Malware." *Electronic Frontier Foundation* 15 March.

Galperin, Eva, Morgan Marquis-Boire and John Scott-Railton. 2013. Quantum of Surveillance: Familiar Actors and Possible False Flags in Syrian Malware Campaigns. Technical report Citizen Lab and Electronic Frontier Foundation.

Garrett, R. Kelly. 2006. "Protest in an Information Society: a review of literature on social movements and new ICTs." *Information, Communication & Society* 9(2):202–224.

Gleditsch, Nils Petter, Jonas Nordkvelle and Håvard Strand. 2014. "Peace research – Just the study of war?" *Journal of Peace Research* 51(2):145–158.

Gohdes, Anita R. 2014. "Predicting unreported conflict fatalities with log-linear mulitple recapture models." Visions in Methodology Conference Paper 2014.

Goldstein, Joshua and Juliana Rotich. 2008. "Digitally networked technology in Kenya's 2007–2008 post-election crisis." *Berkman Center Research Publication* (2008-09).

Goldstein, Robert Justin. 1978. *Political repression in modern America from 1870 to the present*. Cambridge, MA: Schenkman.

González-Bailón, Sandra, Javier Borge-Holthoefer, Alejandro Rivero and Yamir Moreno. 2011. "The Dynamics of Protest Recruitment through an Online Network." *Nature* 1:1–7.

Hafner-Burton, Emilie M, Kiyoteru Tsutsui and John W Meyer. 2008. "International human rights law and the politics of legitimation repressive states and human rights treaties." *International Sociology* 23(1):115–141.

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

Hassanpour, Navid. 2013. "Media Disruption and Revolutionary Unrest: Evidence from Mubarak's Quasi-Experiment." *Political Communication* 30:1–24.

Hathaway, Oona A. 2007. "Why do countries commit to human rights treaties?" *Journal of Conflict Resolution* 51(4):588–621.

Herreros, Francisco and Henar Criado. 2009. "Pre-emptive or Arbitrary." *Journal of Conflict Resolution* 53(3):419–445.

Holliday, Joseph. 2013. "The Assad Regime: From Counterinsurgency To Civil War." *The Institute for the Study of War* .

Hoover Green, Amelia. 2011. Repertoires of Violence Against Noncombatants: The Role of Armed Group Institutions and Ideologies PhD thesis Yale University.

Hounshell, Blake. 2011. "The Revolution Will Be Tweeted." *Foreign Policy* 20 June.

Howard, Philip N. 2010. *The Digital Origins of Dictatorship and Democracy: Information Technology and Political Islam: Information Technology and Political Islam*. Oxford University Press, USA.

Howard, Philip N. 2011. "Data Share: When Do States Disconnect Digital Networks? 1985-2011." Dataset available at: `http://goo.gl/i23FVZ`.

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

Howard, Philip N. and Muzammil M. Hussain. 2013*a*. *Democracy's fourth wave? Digital media and the Arab Spring*. Oxford studies in digital politics Oxford [u.a.]: Oxford Univ. Press.

Howard, Philip N. and Muzammil M Hussain. 2013*b*. *State Power 2.0: authoritarian entrenchment and political engagement worldwide*.

Howard, Philip N., Sheetal D. Agarwal and Muzammil M. Hussain. 2011. "When Do States Disconnect Their Digital Networks? Regime Responses to the Political Uses of Social Media." *The Communication Review* 14(3):216–232.

HRW, Human Rights Watch. 2014*a*. "World Report 2014: Egypt." *Human Rights Watch World Report* .

HRW, Human Rights Watch. 2014*b*. ""They Know Everything We Do": Telecom and Internet Surveillance in Ethiopia." *Human Rights Watch* .

Hultman, Lisa. 2009. "The Power to Hurt in Civil War: The Strategic Aim of RENAMO Violence." *Journal of Southern African Studies* 35(4):821–834.

Humphreys, Macartan and Jeremey M. Weinstein. 2006. "Handling and Manhandling Civilians in Civil War." *American Political Science Review* 100(03):429–447.

Hunziker, Philipp M. and Nils-Christian Bormann. 2013. "Size and Wealth in the International System: Population and GDP per capita Data for Political Science." *Unpublished Working Paper* .

International Working Group for Disease Monitoring and Forecasting. 1995*a*. "Capture- recapture and multiple-record systems estimation I: History and theoretical develop- ment." *American Journal of Epidemiology,* 142:1047–1058.

International Working Group for Disease Monitoring and Forecasting. 1995*b*. "Capture-recapture and multiple-record systems estimation II: Applications in human diseases." *American Journal of Epidemiology* 142:1059–1068.

Jurka, Timothy P., Loren Collingwood, Amber E. Boydstun, Emiliano Grossman and Wouter van Atteveldt. 2012. "RTextTools: Automatic Text Classification via Supervised Learning. R package version 1.3.9." `http://CRAN.R-project.org/package=RTextTools`.

Kalyvas, Stathis. 2006. *The Logic of Violence in Civil War*. Cambridge Univeristy Press.

Kalyvas, Stathis and Laia Balcells. 2010. "International System and Technologies of Rebellion: How the End of the Cold War Shaped Internal Conflict." *American Political Science Review* 104:415–429.

Kalyvas, Stathis and Matthew A. Kocher. 2009. "The dynamics of violence in Vietnam: An analysis of the hamlet evaluation system (HES)." *Journal of Peace Research* 46(3):335–355.

Keating, Joshua. 2013. "Firing Mortars? There's an App for That." *Slate* 18 September.

Keck, Margaret E and Kathryn Sikkink. 1998. *Activists beyond borders: Advocacy networks in international politics*. Cambridge Univ Press.

Kellow, Christine L. and H. Leslie Steeves. 1998. "The role of radio in the Rwandan genocide." *Journal of Communication* 48(3):107–128.

Kern, Holger Lutz and Jens Hainmueller. 2009. "Opium for the Masses: How Foreign Media Can Stabilize Authoritarian Regimes." *Political Analysis* 17(4):377–399.

Kim, HeeMin, Jenifer Whitten-Woodring and Patrick James. 2014. "The Role of Media in the Repression–Protest Nexus: A Game-theoretic Model." *Journal of Conflict Resolution* XX.

King, Gary, Jennifer Pan and Margaret E Roberts. 2013. "How Censorship in China Allows Government Criticism but Silences Collective Expression." *American Political Science Review* 107:1–18.

King, Gary, Jennifer Pan and Margaret E. Roberts. 2014. "Reverse-engineering censorship in China: Randomized experimentation and participant observation." *Science* 345:1–10.

King, Gary and Will Lowe. 2006. "10 million International Dyadic Events." *Dataset [Version 5]* http://gking.harvard.edu/data.

Kocher, Matthew Adam, Thomas B. Pepinsky and Stathis N. Kalyvas. 2011. "Aerial Bombing and Counterinsurgency in the Vietnam War." *American Journal of Political Science* 55(2):201–218.

Krüger, Jule. 2014. "Overlap analysis and certainty in the empirical study of conflict and violence: A comparison of the reporting of civilian casualties in Sierra Leone, 1997- 2001." *Peace Science Society Workshop Paper* .

Krüger, Jule and Christian Davenport. 2014. "Who targets whom, how, and why? A multidimensional approach to understanding violence." *Working Paper* .

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

Landler, Mark and Brian Stelter. 2009. "Washington Taps Into a Potent New Force in Diplomacy." *The New York Times* 16 June.

Landman, Todd and Marco Larizza. 2009. "Inequality and Human Rights: Who Controls What, When, and How." *International Studies Quarterly* 53(3):715–736.

Lesnoff, Matthieu and Renaud Lancelot. 2013. *aods3: analysis of overdispersed data using S3 methods*. aods3 package version 0.4-1.

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

Lichbach, Mark I. 1995. *The rebel's dilemma*. University of Michigan Press.

Little, Andrew. 2014. "Communication Technology and Protest." *Cornell University Working Paper* .

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

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

Lorentzen, Peter L. 2013. "Regularizing rioting: permitting public protest in an authoritarian regime." *International Quarterly Journal of Political Science* 8(2):127–158.

Lotan, Gilad, Erhardt Graeff, Mike Ananny, Devin Gaffney, Ian Pearce and Danah Boyd. 2011. "The Arab Spring| The Revolutions Were Tweeted: Information Flows during the 2011 Tunisian and Egyptian Revolutions." *International Journal of Communication* 5(0):1375–1405.

Lum, Kristian, Megan Emily Price and David Banks. 2013. "Applications of Multiple Systems Estimation in Human Rights Research." *The American Statistician* 67(4):191–200.

Lum, Kristian, Megan Price, Tamy Guberek and Patrick Ball. 2010. "Measuring Elusive Populations with Bayesian Model Averaging for Multiple Systems Estimation: A Case Study on Lethal Violations in Casanare, 1998-2007." *Statistics, Politics, and Policy* 1(1)(1):2–26.

Lyall, Jason. 2009. "Does Indiscriminate Violence Incite Insurgent Attacks?: Evidence from Chechnya." *Journal of Conflict Resolution* 53(3):331–362.

Lyall, Jason. 2010. "Are Coethnics More Effective Counterinsurgents? Evidence from the Second Chechen War." *American Political Science Review* 104(1):1–20.

Lynch, Marc. 2011. "After Egypt: The Limits and Promise of Online Challenges to the Authoritarian Arab State." *Perspectives on Politics* 9:301–310.

Lynch, Marc. 2013. "The Political Science of Syria's War." Project on Middle East Political Science.

Lynch, Marc, Deen Freelon and Sean Aday. 2014. "Blogs and Bullets III: Syria's Social Mediated War." *United States Institute of Peace Peaceworks* 91.

Lyon, David. 2009. *Surveillance studies : an overview*. Repr ed. Cambridge: Cambridge, MA: Polity Press.

MacFarquhar, Neil and Hwaida Saad. 2012. "Rebel Groups in Syria Make Framework for Military." *The New York Times* 7 December.

Mackey, Robert. 2014. "Jailed for Protests, Activists in Egypt and Bahrain Turn to Hunger Strikes." *The New York Times* 4 September.

MacKinnon, Rebecca. 2012. *Consent Of The Networked: The Worldwide Struggle For Internet Freedom*. New York: Basic Books.

Madory, Doug. 2013. "Internet Blackout in Sudan." *Renesys Blog* 25 September.

Manrique-Vallier, Daniel, Megan E. Price and Anita R. Gohdes. 2013. Multiple Systems Estimation Techniques for Estimating Casualties in Armed Conflicts. In *Counting Civilian Casualties: An Introduction to Recording and Estimating Nonmilitary Deaths in Conflict*. New York: Oxford University Press pp. 165–182.

Marczak, William R., John Scott-Railton, Morgan Marquis-Boire and Vern Paxson. 2014. When Governments Hack Opponents: A Look at Actors and Technology. In *Proceedings of the 23rd USENIX Security Symposium*.

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

McCormick, James M. and Neil J. Mitchell. 1997. "Human Rights Violations, Umbrella Concepts, and Empirical Analysis." *World Politics* 49:510–525.

Miller, Elhanan. 2012. "Syrian opposition uses home-made rockets and Google technology, video reveals." *The Times of Israel* 20 August.

Mitchell, Neil J. 2004. *Agents of Atrocity. Leaders, Followers, and the Violation of Human Rights in Civil War*. New York: Palgrave Macmillan.

Mitchell, Neil J. and James M. McCormick. 1988. "Economic and Political Explanations of Human Rights Violations." *World Politics* 40:476–498.

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

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

Morozov, Evgeny. 2012. *The net delusion: The dark side of Internet freedom*. New York: Public Affairs.

Most, Benjamin A and Harvey Starr. 1989. *Inquiry, logic, and international politics*. Columbia: University of South Carolina Press.

Murdie, Amanda M. and David R. Davis. 2012. "Shaming and Blaming: Using Events Data to Assess the Impact of Human Rights INGOs1." *International Studies Quarterly* 56(1):1–16.

Murphy, Heather. 2014. "Ominous Text Message Sent to Protesters in Kiev Sends Chills Around the Internet." *The New York Times* 22 January.

Nowak, Manfred. 1993. *UN covenant on civil and political rights: CCPR commentary*. Kehl: NP Engel Kehl.

OECD. 2011. "The economic impact of shutting down Internet and mobile phone services in Egypt." *OECD Report Egypt* 4 February.

Office of the United Nations High Commissioner for Human Rights. 2014. The right to privacy in the digital age. Technical report Report of the Office of the United Nations High Commissioner for Human Rights.

Olesen, Alexa. 2014. "The Revolution Will Not be Instagrammed." *Foreign Policy* 29 September.

OpenNet Initiative. 2009. "Internet Filtering in Syria." OpenNet Country Profile.

OpenNet Initiative. 2012. "China." OpenNet Country Profile.

Peksen, Dursun. 2009. "Better or worse? The effect of economic sanctions on human rights." *Journal of Peace Research* 46(1):59–77.

Pemstein, Daniel, Stephen A. Meserve and James Melton. 2010. "Democratic Compromise: A Latent Variable Analysis of Ten Measures of Regime Type." *Political Analysis* 18(4):426–449.

Perlroth, Nicole. 2013. "Syria Loses Access to the Internet." *The New York Times Bits Blog* 7 May.

Peterson, Scott. 2012. "Syria's iPhone insurgency makes for smarter rebellion." *The Christian Science Monitor* 1 August.

Pfeifle, Mark. 2009. "A Nobel Peace Prize for Twitter?" *The Christian Science Monitor* 6 July.

Pierskalla, Jan H. 2009. "Protest, Deterrence, and Escalation: The Strategic Calculus of Government Repression." *Journal of Conflict Resolution* 54(1):117–145.

Pierskalla, Jan H. and Florian M. Hollenbach. 2013. "Technology and Collective Action: The Effect of Cell Phone Coverage on Political Violence in Africa." *American Political Science Review* 107(2):207–224.

Pion-Berlin, David and George A. Lopez. 1991. "Of Victims and Executioners: Argentine State Terror, 1975-1979." *International Studies Quarterly* 35(1):63–86.

Poe, Steven C. 2004. The Decision to Repress: An Integrative Theoretical Approach to the Research on Human Rights and Repression. In *Understanding Human Rights Violations: New Systematic Studies*, ed. Sabine Carey and Steven C. Poe. Cambridge Univ Press pp. 16–38.

Poe, Steven C. and C. Neal Tate. 1994. "Repression of Human Rights to Personal Integrity in the 1980s: A Global Analysis." *American Political Science Review* 88(4):853–872.

Poe, Steven C., C. Neal Tate and Linda C. Keith. 1999. "Repression of the Human Right to Personal Integrity Revisited: A Global Cross-National Study Covering the Years 1976–1993." *International Studies Quarterly* 43(2):291–313.

Price, Megan, Anita R. Gohdes and Patrick Ball. 2014. "Updated Statistical Analysis of Documentation of Killings in the Syrian Arab Republic." *Report commissioned by the Office of the UN High Commissioner for Human Rights* .

Raleigh, Clionadh, Andrew Linke, Håvard Hegre and Joakim Karlsen. 2010. "Introducing ACLED: An Armed Conflict Location and Event Dataset." *Journal of Peace Research* 47(5):651–660.

Raoof, Ramy. 2011. "Egypt: how companies help the government spy on activists." *Global Voices Online* .

Rasha, Abdulla. 2011. "The revolution will be tweeted: the story of digital activism in Egypt." *The Cairo Review of Global Affairs* 3:41–50.

Rasler, Karen. 1996. "Concessions, Repression, and Political Protest in the Iranian Revolution." *American Sociological Review* 61(1):pp. 132–152.

Regan, Patrick M. and Errol A. Henderson. 2002. "Democracy, threats and political repression in developing countries: Are democracies internally less violent?" *Third World Quarterly* 23(1):119–136.

Rejali, Darius M. 2011. *Torture and democracy*. 1 ed. Princeton [u.a.]: Princeton, NJ: Princeton University Press.

Reuter, Ora John and David Szakonyi. 2013. "Online Social Media and Political Awareness in Authoritarian Regimes." *British Journal of Political Science* XX:1–23.

Reynolds, Evangeline M. 2014. "Best of All Plausible Worlds? Checking Robustness of Time-Series Cross-Sectional Models with Fictitious Plausible Alternate Treatments." Visions in Methodology Conference Paper 2014.

Rheingold, Howard. 2008. Mobile media and political collective action. In *Handbook of mobile communication studies*. MA: MIT Press pp. 225–241.

Roberts, Margaret E. 2014. "Fear or Friction? How Censorship Slows the Spread of Information in the Digital Age." *Working Paper* .

Rød, Espen Geelmuyden and Nils B. Weidmann. forthcoming. "Empowering Activists or Autocrats? Internet and Authoritarian Survival." *Journal of Peace Research* XX.

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

Sambanis, Nicholas. 2004. "What Is Civil War?" *Journal of Conflict Resolution* 48(6):814–858.

Schnakenberg, Keith E. and Christopher J. Fariss. 2014. "Dynamic Patterns of Human Rights Practices." *Political Science Research and Methods* 2:1–31.

Schofield, Matthew. 2013. "Memories of Stasi color Germans' view of U.S. surveillance programs." *McClatchy, Washington DC* .

Schrodt, P. A. a. B., John and Mohammed Idris. 2014. "Three'sa Charm?: Open Event Data Coding with EL: DIABLO, PETRARCH, and the Open Event Data Alliance." *ISA Annual Meeting Paper* .

Schrodt, Philip A. 2012. "Precedents, Progress, and Prospects in Political Event Data." *International Interactions* 38(4):546–569.

Schutte, Sebastian. 2014. "Violence and Civilian Loyalties: Evidence from Afghanistan." *Working Paper, University of Konstanz* .

Schutte, Sebastian and Nils B. Weidmann. 2011. "Diffusion patterns of violence in civil wars." *Political Geography* 30(3):143 – 152.

Shadmehr, Mehdi and Dan Bernhardt. 2013. "A Theory of State Censorship." *Working Paper* .

Shapiro, Jacob N. and Nils B. Weidmann. forthcoming. "Is the Phone Mightier than the Sword? Cell Phones and Insurgent Violence in Iraq." *International Organization* XX.

Shirky, Clay. 2008. *Here comes everybody: The power of organizing without organizations*. New York : Penguin Press.

Shirky, Clay. 2011. "Political Power of Social Media-Technology, the Public Sphere Sphere, and Political Change, The." *Foreign Affairs* 90:28.

Shirky, Clay. 2014. "Occupy Hong Kong: Macro scale, micro-adaptations." *GitHub: `https://github.com/cshirky/occupyhongkong` .*

Siegel, David A. 2011. "When Does Repression Work? Collective Action in Social Networks." *Journal of Politics* 73(4):993–1010.

Singel, Ryan. 2011. "Lawmaker Calls for Limits on Exporting Net-Spying Tools." *Wired* 11 February.

Slim, Hugo. 2007. *Killing Civilians: method, madness, and morality in war*. Hurst Publishers Ltd.

Stanton, Jessica. 2009. Strategies of Violence and Retraint in Civil Wars PhD thesis Columbia University.

Steele, Abbey. 2009. "Seeking Safety: Avoiding Displacement and Choosing Destinations in Civil Wars." *Journal of Peace Research* 46(3):419–429.

Sullivan, Christopher M. 2012. "Blood in the Village: A Local-Level Investigation of State Massacres." *Conflict Management and Peace Science* 29(4):373–396.

Sullivan, Christopher M., Cyanne E. Loyle and Christian Davenport. 2012. "The Coercive Weight of the Past: Temporal Dependence and the Conflict-Repression Nexus in the Northern Ireland Troubles." *International Interactions* 38(4):426–442.

Sundberg, Ralph and Erik Melander. 2013. "Introducing the UCDP Georeferenced Event Dataset." *Journal of Peace Research* 50(4):523–532.

Surowiecki, James. 2005. *The wisdom of crowds*. New York: Anchor Books.

Themnér, Lotta and Peter Wallensteen. 2013. "Armed Conflicts, 1946–2012." *Journal of Peace Research* 50(4):509–521.

Tufekci, Zeynep. 2014. "The Day the Turkish Government Banned Itself From Twitter." *Medium: Technology and Society* April 2nd.

Tufekci, Zeynep and Christopher Wilson. 2012. "Social Media and the Decision to Participate in Political Protest: Observations From Tahrir Square." *Journal of Communication* 62(2):363–379.

UCDP/PRIO. 2014. "UCDP/PRIO Armed Conflict Dataset Codebook." Version 4-2014.

UN, United Nations General Assembly. 2013. "Resolution 68/167. The right to privacy in the digital age." *Adopted by the General Assembly on 18 December 2013* .

Valentino, Benjamin A., Paul Huth and Dylan Balch-Lindsay. 2004. "Draining the Sea: Mass Killing and Guerrilla Warfare." *International Organization* 58(02):375–407.

Van Belle, Douglas A. 1997. "Press Freedom and the Democratic Peace." *Journal of Peace Research* 34(4):405–414.

van Dijk, Jan A. 2012. *The network society*. 3 ed. London [u.a.]: London: Sage.

Wagner, Ben and Claudio Guarnieri. 2014. "German Companies Are Selling Unlicensed Surveillance Technologies to Human Rights Violators – and Making Millions." *Global Voices Online* .

Warren, T. Camber. 2014. "Not by the Sword Alone: Soft Power, Mass Media, and the Production of State Sovereignty." *International Organization* 68(1):111–141.

Weidmann, Nils B. 2014*a*. "On the Accuracy of Media-based Conflict Event Data." *Journal of Conflict Resolution* XX:1–21.

Weidmann, Nils B. 2014*b*. "The Violence We Do Not See: Reporting Bias in Conflict Event Data." *ISA Annual Meeting Paper 2014* .

Weidmann, Nils B., Doreen Kuse and Kristian Skrede Gleditsch. 2010. "The Geography of the International System: The CShapes Dataset." *International Interactions* 36(1):86–106.

Weidmann, Nils B. and Kristian Skrede Gleditsch. 2010. "Mapping and Measuring Country Shapes." *The R Journal* 2(1):18–24.

Weinstein, Jeremey M. 2007. *Inside rebellion: The politics of insurgent violence*. Cambridge University Press.

Wickham, Hadley. 2009. *ggplot2: elegant graphics for data analysis*. Springer.

Wintrobe, Ronald. 1998. *The Political Economy of Dictatorship*. Cambridge: Cambridge University Press.

Wood, Elisabeth J. 2003. *Insurgent collective action and civil war in El Salvador*. Cambridge University Press.

Wood, Elisabeth J. 2010. Sexual Violence During War: Variation and Accountability. In *Collective Violence and International Criminal Justice*, ed. Alette Smeulers. Vol. 8 Antwerp: Intersentia chapter 13, pp. 297–324.

Wood, Reed M. and Mark Gibney. 2010. "The Political Terror Scale (PTS): A re-introduction and a comparison to CIRI." *Human Rights Quarterly* 32(2):367–400.

York, Jillian and Trevor Timm. 2012. "Satphones, Syria, and Surveillance." EFF Deeplinks Blog 23 February: `https://www.eff.org/deeplinks/2012/02/satphones-syria-and-surveillance`.

Youmans, William L and Jillian C York. 2012. "Social Media and the Activist Toolkit: User Agreements, Corporate Interests, and the Information Infrastructure of Modern Social Movements." *Journal of Communication* 62(2):315–329.

Zanger, Sabine C. 2000. "A Global Analysis of the Effect of Political Regime Changes on Life Integrity Violations, 1977-93." *Journal of Peace Research* 37(2):213–233.

Zeitzoff, Thomas. 2011. "Using Social Media to Measure Conflict Dynamics: An Application to the 2008–2009 Gaza Conflict." *Journal of Conflict Resolution* 55(6):938–969.

Zhukov, Yuri M. 2013. "An Epidemic Model of Violence and Public Support in Civil War." *Conflict Management and Peace Science* 30(1):24–52.

Zuckerman, Ethan. forthcoming. Cute Cats to the Rescue? Participatory Media and Political Expression. In *Youth, New Media and Political Participation*, ed. Danielle Allen and Jennifer Light. Cambridge, MA: MIT Press.

Major disruptions, global sample 1995-2010

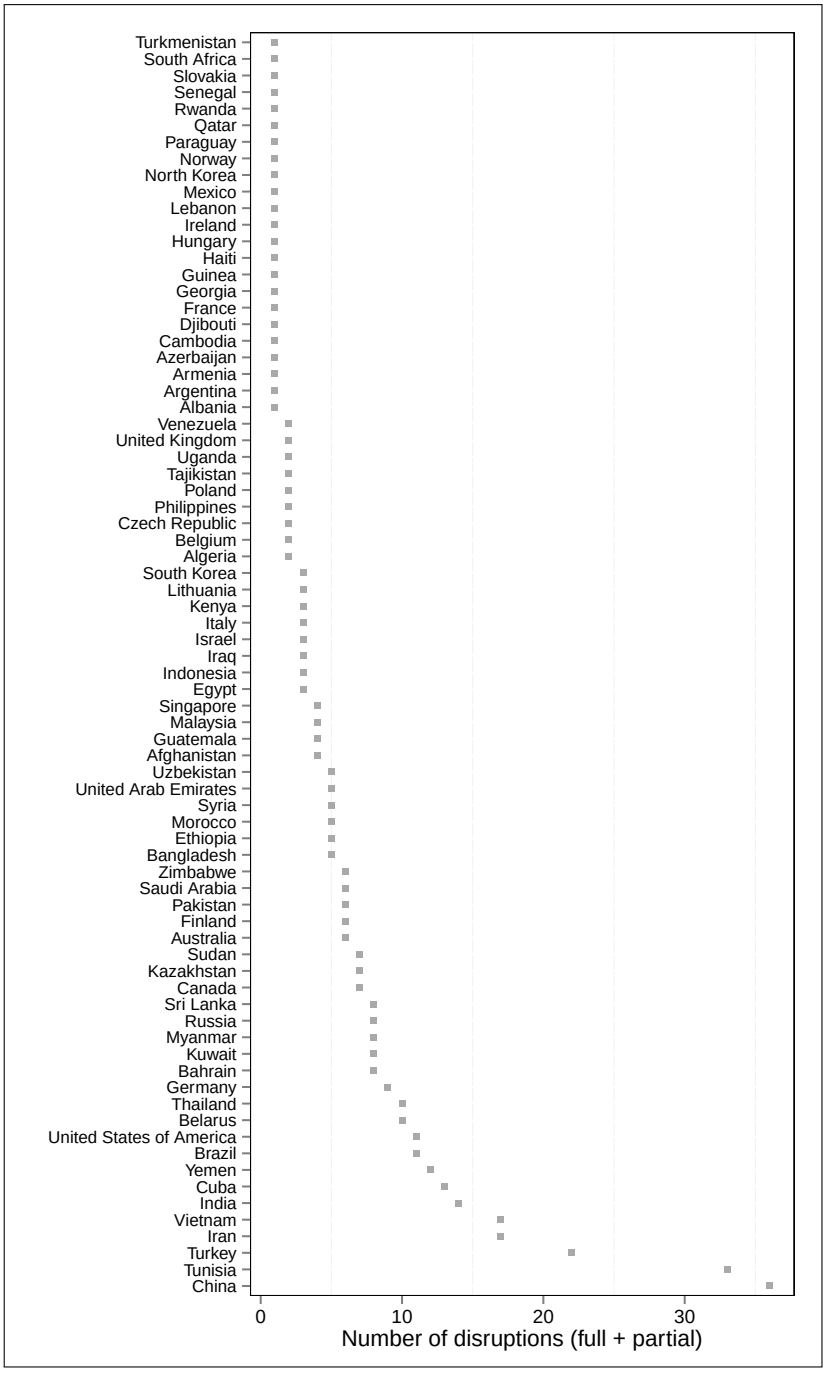


Figure 1. Major government directed internet disruptions (full and partial disruptions).

Network disruptions documented by google traffic

Graphs were created and copied from `https://www.google.com/transparencyreport/traffic/`. All rights belong to Google.

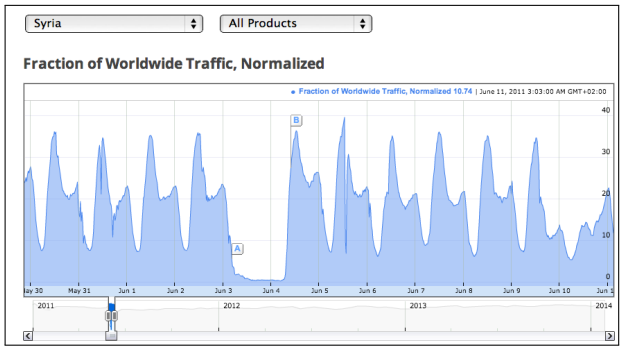


Figure 2. Disrupted Google Traffic, June 2011.

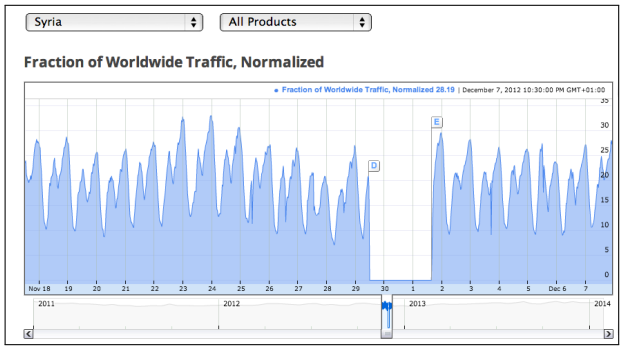


Figure 3. Disrupted Google Traffic, Nov/Dec 2012.

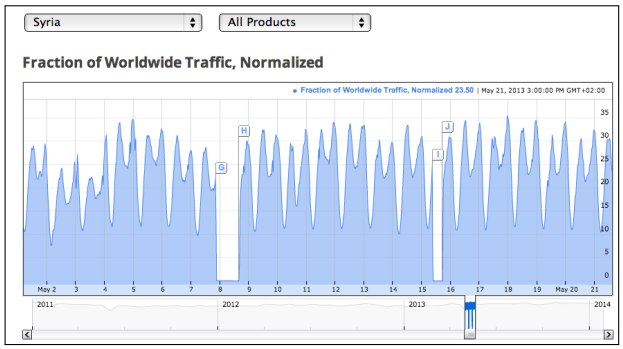


Figure 4. Disrupted Google Traffic, May 2013.

Expected proportion of targeted killings, additional tests

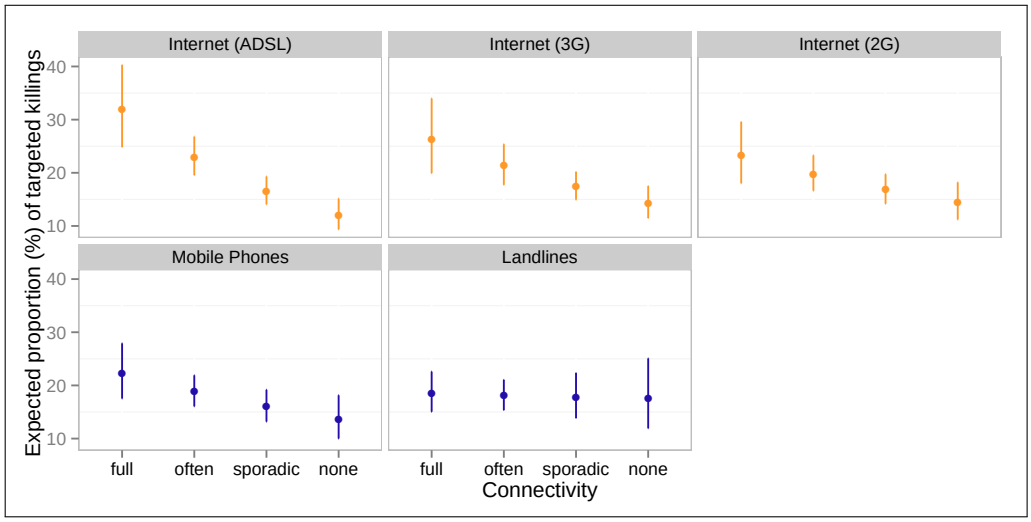


Figure 5. Expected proportion of targeted killings, by Syrian governorate, using the upper bound of estimated killings.

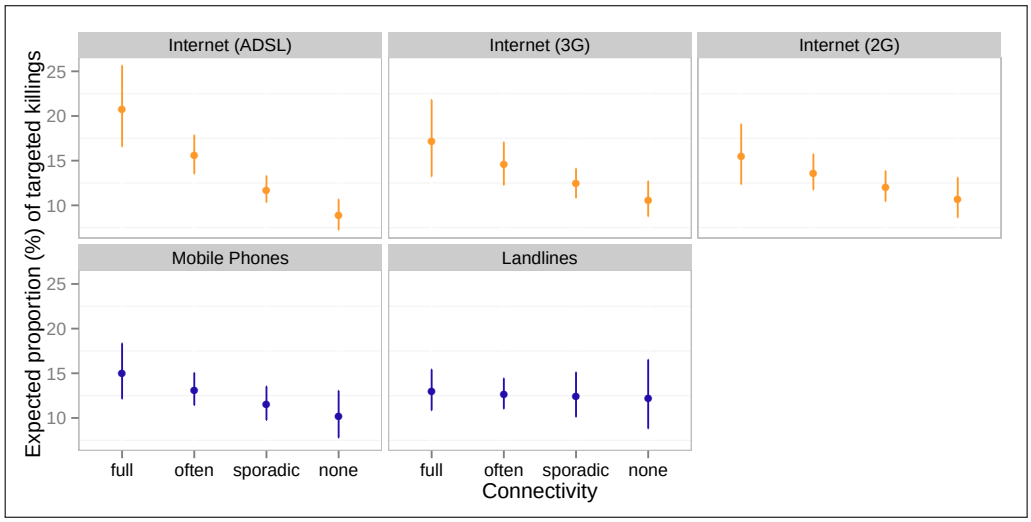


Figure 6. Expected proportion of targeted killings, by Syrian governorate, using the lower bound of estimated killings.