Empirical Advances in the Measurement and Analysis of Violent Conflict

DISSERTATION
zur Erlangung des akademischen Grades
Doctor of Philosophy (Ph.D.)

eingereicht an der
Lebenswissenschaftlichen Fakultät der Humboldt-Universität zu Berlin

von
Ghassan Baliki

Prädentin der Humboldt-Universität zu Berlin
Prof. Dr-Ing. Dr. Sabine Kunst

Dekan der Lebenswissenschaftlichen Fakultät der Humboldt-Universität zu Berlin
Prof. Dr. Bernhard Grimm

GutachterInnen
1. **Prof. Dr. Tilman Brück**
2. **Prof. Dr. Harald Grethe**
3. **Prof. Dr. Nikolaus Wolf**

Tag der müdlichen Prüfung: **05.09.2017**

Abstract

Violent conflict is one of the most persistent challenges affecting the economic livelihoods and food security of individuals worldwide. Despite the surge in literature studying the impacts and drivers of armed conflict, there remains notable knowledge and methodological gaps, particularly regarding the quality of conflict event data. Using various advanced econometric and statistical techniques, this monograph contributes empirically to this literature by studying three interrelated issues. (i) The impact of violence exposure on radicalization; (ii) the magnitude of selection and veracity biases in media-based conflict event data; and (iii) the significance of incorporating violence in nearby locations in predicting armed conflict onset and escalation. First, evidence from the 2009 war on Gaza shows that individuals who experienced violence directly are less likely, on average, to support radical groups. However, when controlling for past electoral preferences, the results reveal a polarization effect among voters exposed directly to violence. Second, by matching conflict event data from several international and national media sources on the Syrian war, media reports are found to capture less than 10% of the estimated total number of events in the study period. Moreover, reported events across the sources exhibit a systematic spatial clustering and actor-specific biases. Third, using a grid-level panel dataset, the temporal and spatial dynamics of violence, among other geographic factors, are found to significantly drive both conflict onset and escalation. However, violence in neighbouring grids does not enhance the prediction of armed conflict when using high precision units of analysis. In addition to these main findings, I propose and discuss a novel methodology, namely crowdseeding, for collecting conflict event data which works directly with primary sources on the ground to provide reliable information for researchers and policy-makers alike.

Keywords: Armed Conflict; Event Data; Violence; Radicalisation; Syria; Gaza.

Abstrakt

Gewaltsamer Konflikt ist eine der hartnäckigsten Bedrohungen des Lebensunterhalts und der Nahrungssicherheit von Individuen weltweit. Trotz einer wachsenden Literatur, die die Ursachen und Folgen von Konflikten untersucht, bestehen nach wie vor erhebliche Verständnislücken, die zum Teil auf einen Mangel an qualitativ hochwertigen Konfliktereignisdaten zurückgehen. Mit Hilfe moderner ökonometrischer und statistischer Methoden trägt diese Monographie empirisch zur Literatur bei, indem sie sich mit drei miteinander verknüpften Themen befasst: (i) die Auswirkungen von Gewalterfahrungen auf Radikalisierung; (ii) das Ausmaß von Verzerrungen („bias“) in medienbasierten Konfliktereignisdaten; sowie (iii) die Rolle von Gewalt in benachbarten Gebieten für die Vorhersage von Ausbruch und Eskalation von Konflikten. Erstens zeigt eine Analyse des Gaza-Krieges von 2009, dass Menschen, die Gewalt direkt ausgesetzt sind, radikale Gruppen im Durchschnitt weniger unterstützen. Wenn frühere Wahlpräferenzen statistisch einbezogen werden, besitzt Gewalt jedoch eine polarisierende Wirkung im Wahlverhalten. Zweitens schätzt eine Auswertung syrischer Konfliktereignisdaten basierend auf internationalen und nationalen Quellen, dass Medien über nur knapp zehn Prozent der auftretenden Ereignisse berichten. Zudem ist die Berichterstattung stark räumlich und nach Konflikt-Akteuren verzerrt. Drittens stellt sich anhand von Paneldaten kleiner geographischer Zellen heraus, dass die räumliche und zeitliche Dynamik von Gewalt starken Einfluss auf sowohl den Ausbruch als auch die Eskalation von Konflikten an einem bestimmten Ort hat. In hochaufgelösten Analysen erhöht Gewalt in benachbarten Raumzellen jedoch nicht die Vorhersagekraft des Modells. Auf Grundlage der empirischen Befunde entwickelt diese Arbeit eine neue Methode zur Erhebung von Konfliktdaten, die auf direkte Informationsquellen vor Ort zurückgreift („crowdseeding“), um Politik und Forschung verlässlichere Daten zu bieten.

Schlüsselwörter: Bewaffneter Konflikt; Ereignisdaten; Gewalt; Radikalisierung; Syrien; Gaza.

Acknowledgements

First of all, I would like to whole-heartedly extend my gratitude to my supervisor, mentor and friend, Tilman Brück. This work would not have been possible without his continuous unconditional support throughout my PhD Journey.

Most of the analyses and results in this dissertation are based on a dataset that I collected in collaboration with Anouk Rigternik and Mary Kaldor at the London School of Economics. I am thankful for their team work and on-going collaboration. I would like to thank the coders for their tireless effort and patience in entering the event data of DISC: Ghiath Agha, Amal Dib, Rasha Natour, Capri McDuffie, and Firas Talhouk. Moreover, I would like thank Tewodros Kebede and Age Tiltnes from fafo in Norway for providing the polling data in Gaza and UNOSAT for sharing their satellite data.

Moreover, I would like to acknowledge the fruitful input, on various presentations of this work, from the fellows and participants of the TAMNEAC project. This EU-funded Marie Curie Network provided me with generous funding to move to Berlin to undertake my doctoral studies. I would also like to thank IGZ for its support to complete this dissertation.

Much gratitude goes to the 2011-2015 team of the Department of Development and Security at DIW Berlin, including: Damir Esenaliev, Kati Schindler, Kerstin Ringelhan, Myroslava Purska, Neil Ferguson, Sindu Kebede, Wolfgang Stojetz, and especially, Veronika Bertram-Hümmer, with whom I really enjoyed writing a paper on wealth and deprivation in Kyrgyzstan. Moreover, a special thanks goes to my friends and colleagues: Florian Szücs, Florian Mölders, and Juliana Werneburg.

Without the encouragement from my family, this journey would have been a much more difficult one. Thank you for believing in me. Most importantly, my sincerest gratitude goes to my life partner, best friend, and the mother of our wonderful daughter Dima. Thank you, Anja Müller-Syring, for your love, support, and patience.

To Dima

Life's sweetness is in its simplicity

Contents

List of Tables . . . XII
List of Figures . . . XIV
List of Abbreviations . . . XIV

Chapter 1. Introduction

1.1 Motivation . . . 2
1.2 Literature Gaps and Contributions . . . 4
1.3 Data and Methodology . . . 6
1.4 Main Findings . . . 9
1.5 Concluding Remarks . . . 11

Chapter 2. The Impact of Violence on Electoral Preferences: Evidence from the Gaza Strip

2.1 Introduction . . . 14
2.2 Literature Review . . . 17
2.2.1 Theoretical Literature . . . 17
2.2.2 Recent Empirical Literature . . . 18
2.3 A Brief Overview of the Gaza War . . . 21
2.4 Data . . . 22
2.4.1 Opinion Polls Data . . . 22
2.4.2 Indices of Violence Exposure . . . 27
2.5 Empirical Framework . . . 32
2.5.1 Identification Strategy . . . 32
2.5.2 Econometric Model . . . 33
2.6 From Violence to Voting . . . 35
2.6.1 Voting Extreme . . . 35
2.6.2 Voting Moderate . . . 41
2.6.3 Results from the Multinomial Approach . . . 43
2.7 Robustness Checks . . . 45
2.7.1 Sample Selection Bias . . . 45
2.7.2 The Undecided Group . . . 48
2.7.3 Political Attitudes as Dependent Variables . . . 49
2.8 Conclusion . . . 52

Chapter 3. Media Bias in Measuring Conflict: Evidence from Syria

3.1 Introduction . . . 56
3.2 Literature Review . . . 60
3.2.1 From Macro to Micro . . . 61
3.2.2 Current State of Micro-level Data . . . 63
3.2.3 Bias in Media Reporting . . . 65
3.3 Context . . . 68
3.4 Data - DISC . . . 71
3.4.1 Design . . . 72
3.4.2 The Coding Process . . . 80
3.4.3 Descriptive Results . . . 81
3.5 Methodological Approach . . . 89
3.5.1 Measuring Selection Bias . . . 89
3.5.2 Measuring Veracity Bias . . . 97
3.6 Main Results . . . 99
3.6.1 Prevalence of Selection Bias . . . 99
3.6.2 Spatial Intensity of the Selection Bias . . . 105
3.6.3 Actor-based Veracity Bias . . . 111
3.6.4 Interpretation of the Results . . . 113
3.7 Crowdseeding: A Novel Approach for Collecting Conflict Event Data . . . 116
3.7.1 Introducing Crowdseeding . . . 117
3.7.2 Type of Events . . . 118
3.7.3 Methodology and Approach . . . 120
3.7.4 Merits and Limitations of Crowdseeding . . . 122
3.8 Conclusion . . . 126

Chapter 4. Predicting Temporal and Spatial Diffusion of Violent Armed Conflict

4.1 Introduction . . . 130
4.2 Literature Review . . . 132
4.3 A Gridded Panel Dataset . . . 136
4.4 Methodology . . . 141
4.4.1 Spatial Autologistic Approach - Conflict Onset . . . 142
4.4.2 Spatial Panel Maximum Likelihood Approach - Conflict Intensity 143
4.5 Main Results . . . 144

4.5.1 Conflict Onset 144
4.5.2 Conflict Intensity 147
4.6 Robustness Checks and Statistical Corrections 153
4.7 Conclusion 157

Bibliography

Appendices

Appendix to Chapter 2 170
Appendix to Chapter 3 174
Appendix to Chapter 4 184

List of Tables

Chapter 2

2.1 Palestinian parties, political position, and share of votes 25
2.2 Support of halting rocket launching into Israel 26
2.3 Support of peace talks with Israel 26
2.4 Differences in individual and household characteristics between direct exposed and control groups 29
2.5 Impact of direct and indirect violence exposure on voting for extreme parties 37
2.6 Impact of direct and indirect violence exposure on voting for moderate parties 42
2.7 Impact of direct and indirect violence on voting extreme- multinomial appraoch 44
2.8 Tobit II model for sample selection: impact of violence on voting for extreme parties 47
2.9 Impact of violence on the support of peace talks 50
2.10 Impact of violence on the support of launching rockets against Israel . . . 51

Chapter 3

3.1 DISC's time frame: variation between report and event dates 77
3.2 Geographic precision levels in DISC 78
3.3 Spatial grid unit correlations between media sources 88
3.4 Event matching by source 99
3.5 Frequency statistics of the matching algorithm 100
3.6 Abundance $\hat{N}$ estimates from various capture-recapture models 102
3.7 Probabilities of the reporting selection bias 105
3.8 Advantages and disadvantages of crowdseeding 123

Chapter 4

4.1 Spatial determinants of conflict onset - baseline model 145
4.2 Spatial determinants of violence intensity - rook model 149
4.3 RMSE and MEA for in- and out- of sample predictions of violent conflict escalation 151
4.4 Determinants of conflict onset - 10 km grid cells 154

Appendices

A.1 Differences in individual and household characteristics between indirect exposed and control groups . . . 170
A.2 Impact of direct and indirect violence on voting moderate - multinomial approach . . . 171
A.3 Impact of direct and indirect damage on election participation . . . 172
A.4 Impact of direct and indirect damage on voting extreme - subsample . . . 173
A.5 Spatial determinants of violence intensity - queen model . . . 184
A.6 ML and two-stage Heckman selection models - determinants of violence escalation . . . 185

List of Figures

Chapter 1

1.1 Number of active violent armed conflicts 1989 - 2015 3

Chapter 2

2.1 Location of damage in the Gaza Strip - 2009 war 22
2.2 Location damage in Gaza - governorate level 31
2.3 Marignal effects on the probability of voting extreme 39
2.4 Various effects of exspoure on the probability of voting extreme 40

Chapter 3

3.1 Type of events in DISC - pie chart 82
3.2 Share of reported events' type by source 83
3.3 Media coverage of violent and political events 84
3.4 Gridded spatial distribution of violence incidence by media source 86
3.5 Histogram of $\hat{N}$ 103
3.6 Inhomogeneous simulations of violent events based on various spatial intensities 106
3.7 L-Ripley envelopes with intesnity based on observed data 108
3.8 L-Ripley envelopes with intesnity based on λ 110
3.9 Propensities of reporting an armed actor conditional on reporting any type of violence 112
3.10 Propensities of reporting an armed actor conditional on reporting violence against civilians 113

Chapter 4

4.1 Spatial density distributions of violence in Syria 137
4.2 Gridded spatial distribution of the determinants of violence 140
4.3 ROC - Receiver Operating Characteristic plots of conflict onset 146
4.4 Out-of-sample forecasts for intensity of violence 152
4.5 ROC plots of conflict onset - 10 km grids 155

Appendices

A.1 Boxplot of pearson residuls of capture-recapture models 183

List of Abbreviations

ACLED	Armed Conflict Location Event Data
AFP	Agence France Press
AP	Associated Press
AUC	Area Under Curve
BBCM	British Broadcasting Cooperation - Monitoring
CRS	Complete Spatial Randomness
DISC	Documentation Initiative of the Syrian Conflict
DIW	Deutsches Institut für Wirtschaftsforschung
GAR	Global Assessment Report
GED	Geo-referenced Event Data
IMR	Inverse Mill's Ratio
ISDC	International Development and Security Center
LSE	London School of Economics and Political Science
MAUP	Modifiable Areal Unit Problem
MCMC	Markov Chain Monte Carlo
MEA	Mean Absolute Error
MENA	Middle East and North Africa
ML	Maximum Likelihood
NIS	New Israeli Sheqel
OLS	Ordinary Least Squares
PPM	Point Process Models
PSU	Primary Sampling Units
RMSE	Root Mean Squared Error
ROC	Reciever Operating Characteristics
SANA	Syrian Arab News Agency
SCAPE	Syria Conflict And Peace Events
SES	Socio-Economic Status
TAMNEAC	Training and Mobility Network for the Economic Analysis of Conflict
UCDP	Uppsala Conflict Data Program

Chapter 1

Introduction

1.1 Motivation

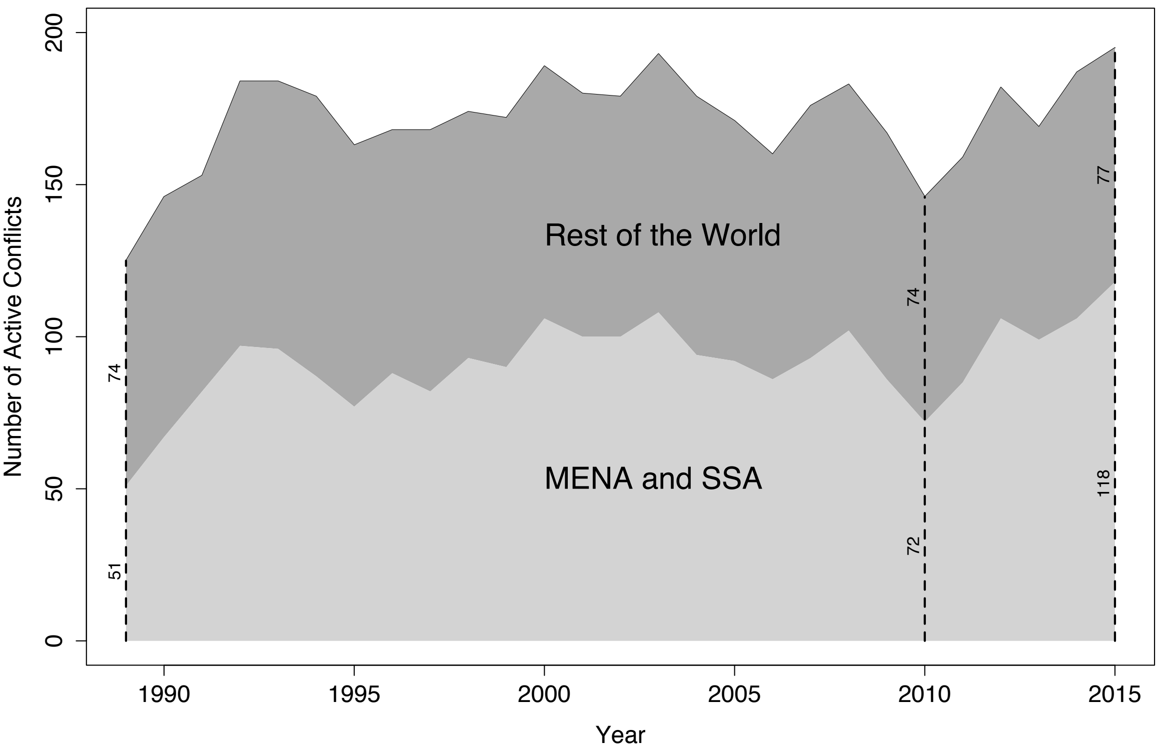
Violent political conflicts cripple economies, wreck public and private infrastructure, and most importantly, result in severe humanitarian crises. The adverse repercussions of violent conflict on the most vulnerable individuals are many-fold. Beyond the devastating loss of life, conflicts force individuals to flee their homes, abandon their lands, and negatively impact their health, food security and socio-economic well-being (e.g., Justino, 2009; Hendrix and Brinkman, 2013; Minoiu and Shemyakina, 2014).

The proportion of countries with large civil wars (i.e., accounting for more than 1,000 battle deaths) decreased significantly since the end of twentieth century (Blattman and Miguel, 2010). However, armed violent conflict remains a widespread phenomena worldwide given that fighting became limited in scale and locally concentrated. In recent years, the rise of violence in the Middle East and North Africa (MENA) after the Arab Spring of 2011 and the continuing spread of armed conflicts in Sub-Saharan Africa (SSA) brought the number of active armed conflicts to an unprecedented level by the year 2015, as shown in figure 1.1. This unexpected rise in the number of armed conflicts prompted researchers to revisit and challenge our understanding of the forms, drivers and dynamics of violent conflict.

The analysis of armed violent conflict developed substantially in the last two decades (Kalyvas, 2005; Cederman, 2008; Verwimp et al., 2009, among others). Research focusing on the socio-economic causes and consequences of violent conflict generated important theoretical and empirical insights on how individuals affected by war behave to secure their livelihoods (Justino et al., 2016). These advances are owed to the surge in data availability at the micro-level collected via household surveys in post-conflict countries, which integrate retrospective self-reported information on war-related experiences alongside a wide range of socio-economic and demographic outcome indicators (Brück et al., 2016).

In order to better study the individual-level impacts of armed violent conflict, it is

Figure 1.1: Number of active violent armed conflicts 1989 - 2015



Note: *Data on active conflicts is extracted from the most recent UCDP-GED dataset and aggregated by year. The lower (upper) area shows the yearly total number of active conflicts in the Middle East and Africa (Rest of the World). The number of active conflicts in the MENA and SSA increased from 72 to 118 between 2010 and 2015.*

imperative to understand the forms, drivers and dynamics of conflict at the micro-level too. Quantitative research of violent conflict benefited greatly from the development of temporally and spatially disaggregated event datasets (Gleditsch et al., 2014). Conflict event data provides geo-coded description on the location, time, actors, and outcomes of violent events at the local level. The Armed Conflict Location and Event Data - ACLED (Raleigh et al., 2010), and the UCDP Geo-referenced Event Dataset - UCDP GED (Sundberg and Melander, 2013) are two widely used sub-national conflict event datasets. These datasets opened the black box of war by enabling the analyses of temporal and geographic variations of warfare (e.g., Buhaug et al., 2009; Weidmann, 2011). Most notably, the disaggregation of armed conflict refocused the attention of researchers from cross-country analyses towards the micro processes of armed conflicts within states.

However, the precision and accuracy of conflict event data remain an issue, particularly regarding its use of media reports as the primary source of information on violence incidence. To date, there are no systematic studies that assess the quality of media-based event data and the effects of reporting biases on the outcomes of the analyses. Furthermore, and beyond the data gaps and challenges, we still have limited rigorous evidence on how armed conflicts start and escalate at the local level, and how the popular support for groups perpetrating violence is sustained.

Against this backdrop, this dissertation has four objectives. First, to investigate how levels of exposure to conflict affect individuals' electoral preferences and their support for armed groups. Second, to undertake a comprehensive quantitative assessment on the quality of conflict event data by determining the extent of the selection and veracity biases resulting from the use of media reports as the main source of information. Third, to propose a novel conflict event data collection methodology, namely crowdseeding, which does not rely on the use of media reports, but instead collects such data from primary sources. Fourth, to study the spatial and temporal dynamics of local armed conflict onset and escalation, and evaluate the usefulness of spatio-temporal models in predicting future incidences of violence.

The rest of this chapter is structured as follows: The next section outlines the main contributions of this monograph based on identified gaps in the literature. Section 1.3 introduces the datasets and the methodologies used for generating the main findings of this thesis, which are summarized in section 1.4. The last section concludes.

1.2 Literature Gaps and Contributions

This dissertation contributes to several interlinked strands of literature in social sciences by proposing new empirical evidence and data methods on the measurement and analysis of armed conflict. Below I present in detail the chapter-specific research gaps and contributions.

First, chapter 2 of this monograph contributes to the growing body of literature which investigates the effects of conflict and violence exposure on the support of radical groups. Studies found consistent evidence linking radicalism and support for extreme political parties to experiences of violent conflict, particularly when the violence is indiscriminate (Lyall, 2009; Gould and Klor, 2010). Most of these studies, however, used indirect measures of exposure to violent conflict such as the proportion of casualties to population size at the district levels (e.g., Jaeger et al., 2012). To my knowledge, there are no studies that investigate the impact of indiscriminate violence on electoral preferences using direct violence exposure indicators. In this chapter, both direct and indirect measures of exposure of violence are applied. The two indicators respectively capture the extent of the damage resulting from the war on households' own homes and other dwellings in their neighbourhood. The novelty of this approach is that it permits a clear distinction between individuals who experienced violence first-hand, second-hand, and others who were not exposed to violence at all.

Second, chapter 3 generates valuable insights to the field of conflict and peace research by testing the severity of the biases in conflict event data. As mentioned above, the surge in the number of studies on the micro-level drivers and consequences of violence is owed to the availability of conflict event data like ACLED and UCDP-GED. These datasets rely heavily on news media reports to gather information on conflict incidence at the local level. Previous work showed that media reports are seldom unbiased in their selectivity of news and their description of the events (Earl et al., 2004; Davenport, 2009, among others). To date, there are only few studies that test the reliability in the use of media reports in providing accurate and unbiased conflict event data (Weidmann, 2016; Zhukov and Baum, 2016). However, there has been no research undertaken that measures precisely the extent of the selection and veracity biases on conflict event data or their spatial accuracy. Based on these gaps, chapter 3 evaluates the reliability of news reports in providing unbiased information by quantitatively and rigorously addressing both types of biases in the data. These contributions echo the urgent need for testing the validity of media-based conflict event data.

Third, the remaining part of chapter 3 offers new pathways in the collection of conflict event data at the local level. Based on a current project in Syria, I propose a novel methodology for the collection of event data which circumvents the need to rely on secondary media sources. Crowdseeding works with primary sources on the ground to generate conflict event data remotely and on a day-to-day basis. The main advantage of this technique is that it mitigates problems arising from the use of media reports, but retains the scientific codebook-oriented classifications of violent events. This methodological contribution of this technique to existing data collection efforts of conflict event data benefits both practitioners and researchers alike.

Lastly, Chapter 4 of this dissertation contributes to the recent literature which attempts to predict the future outbreaks of armed conflict (see Cederman and Weidmann (2017) for a compact review). Improved data accessibility and enhanced modelling have enabled researchers to explore better the spatial and temporal dynamics of armed violence. However, the development of realistic and robust predictive models that enable forecasting outbreaks of conflict are scarce. Moreover, there are no substantial works to date that predict levels of violence escalation. This inadequacy is owed to the difficulty in obtaining sound out-of-sample predictions, which require large and precise data. Chapter 4 provides additional insights in the evaluation of the spatial and temporal predictive capabilities of conflict. This is achieved by analysing the role of space in generating forecasts not just for conflict onset, but also for violence escalation.

1.3 Data and Methodology

Data

This dissertation is based on three distinct datasets, which utilize both types of micro-level conflict data (event data and household surveys).

Chapter 2 uses a representative household survey and individual polling data from Palestine. The survey was collected by fafo Institute of Norway as part of a series of opinion polls conducted between 2005 and 2010. The survey wave used in this dissertation was undertaken in March 2009, less than two months after the end of the Gaza war. The questionnaire included comprehensive details on dwelling and neighbourhood damage, food insecurity, electoral preferences, and political opinions of war and the Israeli-Palestine peace talks. Both damage of household's own dwelling and other dwelling in their vicinity are used as indicators of exposure to violence. The main strength of using such dataset, as discussed earlier, is that is allows a clear identification of individuals experiencing violence. The weakness, however, is that the use of only one cross-section (instead of a panel) diminishes the feasibility of analysing the long-term impacts of violence exposure.

The data used in chapter 3 for measuring media biases in conflict event data was collected as part of my doctoral studies. Conflict event data was generated from Syria using a number of renowned news sources (which included Agence France Press, Associated Press, and BBC Monitoring, among others), for a pre-selected sample time frame. The data collection effort was a collaboration with colleagues at the London School of Economics (LSE). Coders were trained to review the full reports from the list of selected media outlets and record the event data based on a codebook developed solely for this purpose. The dataset included information on the geo-location, dates, actors, and outcomes of violent actions, as well as political actions.

Based on the experiences from this data collection effort, a new dataset was also collected on the Syrian war using the crowdseeding approach. The novelty of the technique is demonstrated through its work with on-the-ground reporters instead of relying on media reports. The methodology behind the dataset is presented in detail in chapter 3 of this dissertation.

Context

The Israeli-Palestinian and Syrian wars are two notable examples of the contemporary forms of violent conflicts. Since 2009, Israel has launched several attacks on the Gaza Strip that resulted in a large number of casualties and a devastating destruction of dwellings. Moreover, given the political and economic isolation of the Gaza Strip, information coming from the region sheds important light on the possible political repercussions of violence. The Syrian war, which broke out in March 2011, is to date one of the most severe humanitarian crises of the 21st century. The severity and changing dynamics of the civil conflict makes it a suitable case for studying temporal and spatial dynamics of armed violent conflict.

Methodologies

All chapters of this dissertation employ advanced quantitative statistical approaches from various disciplines, including economics, political science, geography, and ecology.

In chapter 2, logistic binomial and multinomial estimations are used extensively to tease out the impacts of violence exposure on electoral preferences and political attitudes in the Gaza Strip. These estimations are applied due to the binary and categorical nature of the dependent variables of interest. Individual electoral preferences are primarily transformed into binary variables for both voting for extreme or moderate parties.

Chapter 3 employs capture-recapture techniques, which are widely used in ecological sciences, to determine the magnitude and extent of the underreporting bias across media sources. Capture-recapture approaches apply both loglinear and Markov Chain Monte Carlo (MCMC) simulations to estimate the abundance size of the number of unobserved events based on matched observed data. This approach is useful in the absence of reliable reference sources. Moreover, simulations drawn from Poisson Point Models (PPMs) are used to determine the spatial clustering and dispersion of the reported events for each source.

Chapter 4 investigates the spatial and temporal geographic determinants of armed conflict onset and escalation by using logistic regressions with spatial dependence and panel fixed effects maximum likelihood estimations, respectively. The dataset is built using the conflict event data from Syria which is aggregated to the grid level. Gridded systems exhibit an apolitical boundary of a given country, which are advantageous over using administrative and political boundaries. In-sample and out-of-sample predictions are generated via large iterative MCMC simulations based on Gibbs sampling techniques.

All methodological approaches used for generating the findings in this dissertation are performed using R statistical programme. These advanced econometric and statistical techniques are either found in various existing R packages or are developed solely for the analysis of this dissertation. Alternative approaches and statistical checks are conducted to ensure the robustness of the results.

1.4 Main Findings

In this section, brief summary of the main findings is presented for each chapter of this thesis.

In chapter 2, the main results indicate that individuals who were exposed to direct major damages during the Gaza war are less likely, on average, to vote for extreme parties. However, the effect disappears after controlling for prior electoral preferences, implying that there is no impact on changes in support for extreme parties. Instead, exposure to direct violence exhibits a polarization effect, dividing the political support for extreme and moderate parties in Gaza further apart. In contrast, exposure to indirect violence, measured through reported damage in the immediate neighbourhood, has no significant effect on electoral choices. The outcomes are consistently robust to a number of methodological checks, including controlling for self-selection and the inclusion of absentees and undecided voters in the analysis. The results of chapter 2 draw a contrasting picture from the one already established in the literature (which use only indirect measures), where the short-term impacts of exposure to violence are found to have a clear radicalisation effect (Jaeger et al., 2012).

Chapter 3 has three main findings. First, all media sources in the dataset are found to report less than 10% of the estimated total number of events that took place within the sample time frame. The results are robust to various model specifications and are significant at the 5% level. Second, violence events reported by most of the media sources are highly clustered *vis-á-vis* the estimated expected distribution of the unobserved events. This implies that media reports focus on specific locations in their coverage of the conflict. Third, local news sources exhibit a significant actor-specific bias in their description of violent events. These variations become more evident when media sources report on indiscriminate violence targeted against civilians.

The results from the main findings of chapter 3 reveal a large bias in the number of reported events, their location, and their description of the involved armed actors. These worrisome insights underscore the need for alternative non-media-based conflict event data. In this direction, I introduce and discuss a novel dataset based on a pilot project in Syria, which relies on primary reporters rather than secondary media sources. I discuss the advantages and disadvantages of the various data collection techniques, and conclude that the reliance on primary sources provide researchers with valuable information at the local level and reduces immensely biases plaguing media-based datasets. However, these merits need to be weighed against the rising costs and the safety of the reporters working in such "hot" conflict zones.

Chapter 4 generates two important insights on the determinants and predictions of both violence onset and escalation. First, densely populated and more accessible areas are found to positively affect the onset of violence, which is a finding in line with prior literature. Moreover, the onset of conflict is strongly driven by incidents of violence in previous periods, as well as violence events in neighbouring grids. However, the significance of these determinants are not translated into the predictive capabilities of the model. On the contrary, temporal models are found to perform better than their spatio-temporal counterparts, which are consistent under various specifications and robustness checks, particularly for the out-of-sample predictions. The findings contrast prior literature, which concluded that violence in neighbouring regions plays a vital role in the onset of armed conflict (Ward and Gleditsch, 2002; Weidmann and Ward, 2010).

In regards to violence escalation, the results show that the intensity of violence in a given location de-escalates when there is an increase in violence in neighbouring locations. The forecasting capabilities of the spatio-temporal models of violence intensity are also found to be weaker than temporal ones. The incorporation of variables capturing the incidence of violence in neighbouring regions does not better predict future levels of violence. These results are robust to various specification (such as grid cell size) and are likely to be driven by the clustering of violence incidents recorded from media-based event data. Therefore, the provision of spatially precise event data is an important step towards the enhancement of the predictive capabilities of violence escalation.

1.5 Concluding Remarks

This dissertation highlights and provides evidence on some important gaps in the measurement and analysis of violent conflict. First of all, indicators of violence exposure need to be further assessed to ensure that they indeed capture what we originally intend them to do. Indices of direct exposure to violence are shown to produce better estimates than indirect ones. Household survey data should integrate modules of conflict to guarantee the development of robust measures of exposure to armed conflict at the individual level. In the absence of this information, results are likely to lead to misleading interpretations.

Second, there is an urgent need to move beyond the use of media-based conflict event data. Media reports on armed conflicts are shown to be plagued with biases, both in terms of its selection of the news and its description. Moreover, media reports are found to not just underreport the actual incidences of violence, but also systematically misrepresent the spatial distribution of the events. This in turn can bias the outcomes of the analyses and the validity of the inferences, especially when violence incidents are matched with household surveys at the micro-level. Motivated by these concerns, this dissertation introduces a novel alternative methodology to mitigate such biases. Crowdseeding conflict event data presents a crucial stepping stone in the correct direction. The technique is still in its infancy, and there is still a need to consistently validate the quality of the data it generates.

Third, predicting future incidents of violent conflict remains a challenging task for conflict and peace researchers. This is especially true in forecasting violence escalation. Nevertheless, with the provision of detailed temporal and spatial data at precise local levels, as well as the application of appropriate methodological approaches, the production of policy-relevant forecasts of violence is surely feasible.

To that end, this dissertation generates vital advances in the measurement of violent conflict at the micro-level, as well as new insights on its forms and drivers. Yet there remain important knowledge and methodological gaps in our understanding of violent conflict, particularly in the absence of reliable precise event data. By addressing these gaps, we are better placed to study the socio-economic effects of war on households living in conflict-affected and fragile settings, and subsequently provide sound policy prescriptions to diminish their adverse impact.

Chapter 2

The Impact of Violence on Electoral Preferences: Evidence from the Gaza Strip

2.1 Introduction

The battle in winning hearts and minds remains a myriad challenge to governments world-wide in their fight against extremism. Insurgents and extreme groups are more likely to consolidate their political power if they are backed by the local population. The use of violence as a counter-insurgency tool is a double-edged sword. On the one hand, the use of violence against radical groups has shown to increase their popular support, especially when the violence is indiscriminate and causes a high number of civilian casualties (Kocher et al., 2011; Lyall, 2009). On the other hand, targeted decapitation of group leaders is found to deter public support to the group in the long-run (Johnston and Sarbahi, 2016). This conundrum is further complicated with the recent globalisation of insurgent violence. Hence, violent counter-insurgency actions may not be a sufficient in curtailing a group's popular support, if not achieving the opposite. Such a challenge is especially heightened when the state is considered an aggressor rather than a liberator, as in the case of the Israeli-Palestinian conflict. The Israeli government has launched several wars on the Gaza Strip in order to topple down the Hamas government. Yet, the public support to the group in the Strip did not seem to diminish over time, raising important questions on the effectiveness in the use violence as a de-radicalisation strategy.

A number of empirical studies have explored various angles of this reasoning, including a handful of studies focusing on the Palestinian-Israeli case (Berrebi and Klor, 2006, 2008; Gould and Klor, 2010; Caruso and Klor, 2012; Jaeger et al., 2012). These studies address the impact of violence on voting preferences for granting political concessions in both Israeli and Palestinian communities. The results do not suggest a simple one-directional impact in the use of violence. For example, Jaeger et al. (2012) find that exposure to violence discourages Palestinians from supporting moderate political groups in the short-term, shifting their preferences towards radical groups. Yet they also find that this shift disappears after 90 days. Moreover, Gould and Klor (2010) show that local insurgent attacks against Israelis induces them to vote for right-wing parties, but at the same time shifting the political landscape in Israel towards more left-leaning policies (such as willingness to grant territorial concessions to Palestinians). These empirical examples show that exposure to violence might increase, decrease, or have no effect in the long-run on the popular support for extreme parties.

Hence, there remains a clear gap in solving the following paradox: Is violence an effective tool in reducing radicalism and/or popular support of radical groups? Or does violence eventually lead to more violence in the future, generating a vicious loop? In order to answer this set of questions, this chapter analyses the impact of violence on the political preferences and attitudes in the Gaza Strip directly after the end of the 2009 Gaza war. By using a novel representative individual-level data of political opinions and attitudes, the chapter examines the immediate impact of aerial bombing (captured via destruction of dwellings) on the electoral support of extreme groups and political opinion on the continued use of violence.

The Israeli-Palestinian conflict is arguably one the most long-lasting contemporary violent conflicts today. Over a period of 60 years, the intensity and dynamic of the conflict has varied considerably. To date, Israel occupies and controls most of the land in the West Bank and is still issuing permits for building new settlements. The expansion of the settlements in the occupied territories of the West Bank is considered one of the main hurdles against a permanent two-state solution. The Gaza Strip on the other hand, which is *de facto* outside the jurisdiction of the Palestinian Authorities, is burdened economically and socially due to the on-going Israeli blockade. Since 2009, Israel has launched three intense wars, albeit short, which have devastated the infrastructure and the economy of the Gaza Strip further. The motive behind these wars was to curb the rise of radicalism in the Gaza Strip, embodied through Hamas. Yet no evidence suggest that the use of violence has been an effective tool in mitigating the political or military power of Hamas. The main reason behind Israel's failure to oust the Hamas "government" in the Gaza Strip could be due to the popular support the group enjoys. Hence focusing the analysis of this study on the 2009 Gaza War is suitable to understanding the pathways through which violence affects popular support to radical groups.

The results can be summarized as follows: Households directly impacted by intense violence are on average less supportive of extreme parties *vis-á-vis* moderate ones. The results contradict previous studies that show that violence has a radicalization effect in the short-term (Jaeger et al., 2012). Nevertheless, the effect disappears when prior political voting patterns are controlled for. This suggests that direct exposure to violence has a polarization effect on the electorate rather than a one-sided de-radicalisation effect. On the other hand, indirect exposure to violence (measured through the destruction of the neighbourhood) has no significant impact on voting behaviour.

In regards to political attitudes, exposure to violence does not have a significant impact on the support of peace talks. Yet these individuals are less likely to support continued use of violence against Israel. Such an outcome is not surprising given the high correlation between voting for moderate parties and opposing the use of violence.[1]

This study builds on the existing literature in three novel aspects. First, the study uses a unique representative dataset on political opinion and attitudes in the Gaza strip right after the end of the war. The dataset provides rich and valuable information on households affected by the war, facilitating the identification of individuals impacted by indiscriminate violence. Second, the study uses a direct measure of exposure to violence - destruction of own dwelling. Previous work on the topic relied mostly on indirect indicators such as the ratio of fatalities per population in a pre-defined geographic location. The use of such indirect proxies restricts an accurate depiction of actual violence exposure levels. Lastly, the dataset focuses solely on the Gaza Strip, which to date has not been studied fully and remains a crucial area study gap. Given the unique role Gaza has on the outcomes of any peace process within the Israeli-Palestinian conflict, results emerging from this study area is paramount for generating sound policy prescriptions.

[1] *It is important to note that given the inability to examine the long-term effects on both political preferences and attitudes, the interpretation of the results should be done with deliberation.*

The rest of the chapter is constructed as follows: The next section provides a literature review on both theoretical and empirical works on the effect of violence on political preferences. Section 2.3 summarizes the 2009 Gaza war. Section 2.4 describes the dataset and the construction of the violence indices, while section 2.5 sets the methodological framework. The main results and their respective robustness checks are presented in sections 2.6 and 2.7, respectively. The last section concludes.

2.2 Literature Review

2.2.1 Theoretical Literature

There is a growing theoretical literature modelling the role of violence as a tool in winning hearts and minds (Rosendorff and Sandler, 2004; Kalyvas et al., 2006; Bueno de Mesquita and Dickson, 2007, among others). Yet there remains no clear consensus regarding the direction of this impact. The first argument states that violence exerted by a state or other armed actors through targeted killing of smaller factions, opposition leaders, or civilians, will inevitably have a counter-effect. In other words, violence will lead to more popular support to insurgent groups. Military measures targeting indiscriminately civilians and armed actors will not just then foster grievances among the population, but might lead them to support or join radicalised groups (Wood, 2003). This boomerang effect of exercising violence as a means of deterrence might lead to vengeance instead, and therefore the cyclical continuation of the use of violence manifests itself through the prolongation of conflicts and leads eventually to a zero-sum game (Schelling, 1980).

The other school of thought argues that violence remains an effective measure in quelling extremism (Ganor, 2011), and hence counterinsurgency measures, including targeted killings, will incapacitate the activity of armed factions and consequently deter their popular support (Yared and Padró i Miquel, 2012). Jaeger and Paserman (2008) model the cyclical nature of violence by differentiating between three responsive actions: deterrence, vengeance, and incapacitation. Nevertheless, a clear-cut popular response to violence exposure in supporting insurgent or terrorist groups depends on plenty of local-specific factors, such as the ethnic composition of the population and the need of protection (Kalyvas et al., 2006). Bueno de Mesquita and Dickson (2007) model theoretically the impact of counter-insurgency on violence to examine if it leads to modernization or radicalisation of the affected population. They argue that radicalisation is more likely to take place when violence is indiscriminate.

2.2.2 Recent Empirical Literature

Empirically, measuring the impact of violence in winning hearts and minds remains difficult to disentangle given the complexity in building appropriate violence exposure indices and good identification strategies. There have recently been numerous attempts to examine how the use of violence might impact the popular political preferences using individual micro-level data. Two strands of literature are relevant in that regard. The political science literature differentiating outcomes of targeted and indiscriminate violence is of vital importance. Different types of violence in counter-insurgencies generate different responses, particularly regarding their impact on political support of radical groups. The second relevant group of studies examines the effects of violent conflict on political participation, activism, and attitudes using household level data.

Indiscriminate Violence

Aerial bombardment has been used consistently to exert superiority in the battle against insurgencies. Although the negative economic impact of bombing has been well established in the literature (Bellows and Miguel, 2006; Brakman et al., 2004), the effects of indiscriminate violence *vis-á-vis* targeted bombardment in quelling insurgency and terrorism remains largely unknown. The literature mainly differentiates between indiscriminate and systematic violence, where the former usually results in higher number of civilian causalities. Indiscriminate bombardment has been shown to empower rebel groups through increased support among the population (Kocher et al., 2011). Yet these results do not necessarily hold true when the intensity of destruction or the number of killings spur (Lyall, 2009).

The issue also extends to the effectiveness of systematic targeted killing, which aims to decapitate the leadership of the rebel armed groups or insurgents (Jaeger et al., 2009; Johnston, 2012; Price, 2012). Johnston and Sarbahi (2016) find that drone attacks targeting head leaders of Al-Qa'ida in Pakistan are an effective tool in curbing insurgencies. Howoever, it is widely believed that targeted killings induce anger among populations, facilitating recruitment and further violence in the future (Pape, 2014; Jordan, 2009).

Identifying the type of violence used by Israel during the Gaza war can be a daunting task. Given that the Israeli army has relied heavily on air power during the Gaza war, indiscriminate violence was inevitable. Although Israel claims that it only systematically targeted Hamas fighters, the high number of civilian casualties proves otherwise. Literature shows that indiscriminate violence is likely to increase power of and support to rebel groups and insurgents (Kocher et al., 2011), yet most studies do not thoroughly examine this link at the individual level. There remains a notable gap in literature linking indiscriminate violence to individual-level electoral preferences, which can provide a better picture on the actual support to radical or rebel groups.

Violence and Political Preferences

Recent studies have used household micro-level conflict data in war-affected countries to disentangle the effects of violent conflict on political participation and preferences. Bellows and Miguel (2009) use the ACLED dataset in Sierra Leone to analyse the post-conflict effects of violence on political participation and collective action. They find that members of households who directly experienced violent conflict are more likely to attend community meetings and vote. Moreover, Blattman (2009) analyses the influence of exposure to violence of ex-combatants in Northern Uganda on voting and political activism, and finds that violence leads to more activism and political participation.

There is also a long-standing line of literature work focusing on the political impacts of the Israeli-Palestinian conflict on both the Palestinian and Israeli populations. (Berrebi and Klor, 2006, 2008; Gould and Klor, 2010; Caruso and Klor, 2012; Jaeger et al., 2012). These studies rely on indirect indicators to measure violence intensity which include, for example, the ratio of fatalities over district population (Jaeger et al., 2012; Brück et al., 2014), number of border closures (Miaari and Sauer, 2011), threat of rockets (Getmansky and Zeitzoff, 2014). To the best of my knowledge, no study yet has used direct measures of violence exposure at the individual level, or has integrated high-precision spatial measures in their estimation. Direct measures of violence exposure are usually integrated in household questionnaire to facilitate the identification of conflict exposure at the individual level (Verwimp et al., 2009). Researchers rely on the subjective assessment of individuals to determine the intensity of violence subjected to households. The questionnaire used in this study includes detailed information on the destruction of own dwelling, as well as the intensity of damage of dwellings in the respondent's immediate neighbourhood. This enables the direct identification of affected households (see section 2.4 for more details on the construction of the violence indices). Moreover, damage inflicted on civilian households through aerial bombardment and long-range shelling captures the extent of indiscriminate violence, which has not been used to identify its real impact of political and electoral preferences.

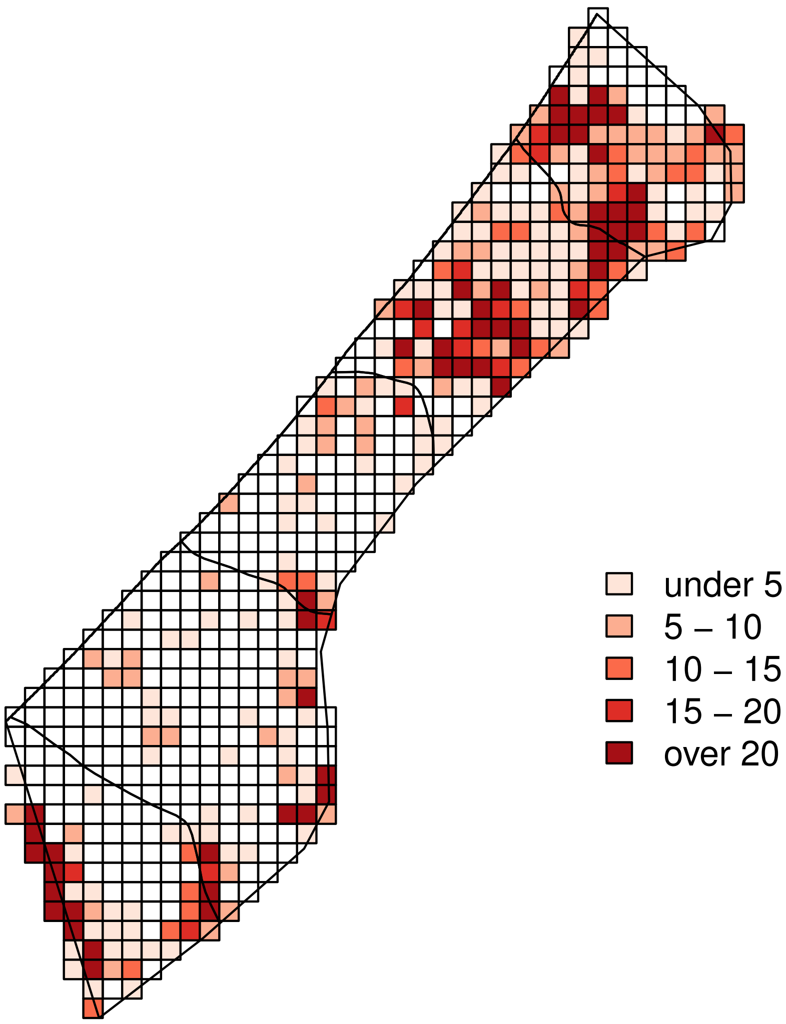
The most relevant work to this study is based on a three-month repeated opinion polls in both the Gaza Strip and the West Bank (Jaeger et al., 2012). In their study, Jaeger et al. (2012) analyse the effect of the ratio of fatalities per district population on the political preference after the second Intifada. The outcome variables are based on the support of peace negotiations and electoral preferences of Palestinian political parties. They found that support for moderate political parties decreases in the short-run, yet this impact diminishes after 90 days. Still, major political events in the Israeli-Palestinian conflict are found to have a long-term impact of political positions. This study builds on the works of (Jaeger et al., 2012), but particularly analyses the immediate impact of the Gaza war of 2009 on the political preferences of civilians affected directly from the war, and not just through district fatality levels. Also the outcome variables used do not just project current support to political factions in the Gaza Strip, but specifically address actual election choices in 2006. Such addition facilitates measuring possible changes in electoral choices between the exposed and the non-exposed groups.

2.3 A Brief Overview of the Gaza War

The Gaza war, also known as operation Cast Lead in Israel, was a three-week intense war of Israel against Hamas in the Gaza Strip during December 2008 and January 2009. The main direct trigger of the war was to stop Palestinian factions in the Gaza Strip (mainly Hamas) from firing rockets into Israel. The assault was full-scale, but the Israelis relied mostly on air power and long-range artillery in their attacks. The tactic was to reiterate against the bases of the rocket lunching inside Gaza. Although the attacks aimed to only target military bases and Hamas combatants, civilians paid a high price for this war. The most inflicted damages, and the highest number of civilian casualties were on the Palestinian side. The war ended in January 2009, without any military or political victories for either side, but with high economic and social costs. The estimated number of deaths resulting from this three-week war was 1,417 on the Palestinian side, out of which 926 were civilians; and only 13 casualties on the Israeli side. Moreover, about 50,800 Gaza residents have been temporarily displaced, and over 4,000 homes were completely destroyed. The approximate cost of damage to Gaza is estimated to be around 2 billion dollars (Zanotti, 2010). Figure 2.1 shows the intensity of the destruction in the Gaza Strip after the end of the attacks. The data was compiled by UNOSAT through using satellite imagery. It includes destruction of dwellings, as well as craters formed in open fields from bombardment.[2] A gridded presentation of the destruction depicts a precise geographical location of the intensity of the damage. Each grid has an area of 0.64 km^2 and contains the aggregate count of destruction points. Most of the destruction is mainly concentrated in residential and dense areas around the northern and southern borders with Israel, but high intensity damage is particularly evident in and around the capital city of Gaza.

[2] *UNOSAT is a technology-intensive programme of UNITAR delivering imagery analysis and satellite solutions to relief and development organisations within and outside the UN system.*

Figure 2.1: Location of damage in the Gaza Strip - 2009 war



Note: *Location data points of the damage was obtained from UNOSAT, the damage includes all recorded bombing sites captured from Satellite imagery.*

2.4 Data

2.4.1 Opinion Polls Data

The data used for this study was collected by Fafo institute, and is based on repeated cross-sectional polls from 2005 till 2010. Due to a number of contextual and methodological considerations, only the 2009 survey is used. First, the 2009 questionnaire aimed specifically to capture the political and electoral preferences after the Gaza war. The survey addressed unique questions related to the war and the destruction which are imperative to the analysis. Second, the polls did not track the same individuals over time, which hinders a proper identification strategy and a sound multivariate panel analysis. Third, the absence of precise location coordinates of households in the datasets restricts the geographic matching of damage-affected areas with previous polls. Hence, based on these concerns and the lack of alternative data sources, the use of one cross-sectional survey is the most appropriate choice despite its analytical disadvantages.

The survey was conducted in February 2009, only one month after the end of the Gaza war. The dataset is representative for the Gaza Strip. The sample frame was designed in accordance with the 1997 census. 132 randomly selected PSUs were used in the first stage of the sampling process. The strata was designed by dividing each governorate proportional to its population into urban, rural, camp areas.[3]

The survey contains two type of questionnaires: the main Household (HH) questionnaire and the randomly selection Individual (RSI) questionnaire. The HH questionnaire addressed war-specific questions on damage and food insecurity and was answered by the main adult household member (usually the household head). The RSI questionnaire contained questions on political opinions, attitudes, and electoral preferences. An eligible member (aged 18 or above) was chosen randomly from the household roster to take part in RSI questionnaire.

Three main survey questions are of particular relevance to this paper:

(i) *"Whom would you vote for if there is an election today?"*[4]

(ii) *"To what extent do you agree with following statement: All Palestinian factions*

[3] *The Gaza Strip constitutes of 5 governorates: North Gaza, Gaza (City), Deir El-Balah, Khan Yunis, and Rafah.*

[4] *Consequently two questions are relevant "Did you participate the 2006 Elections?" and "Whom did you vote for?"*

must stop firing rockets at Israel" and,

(iii) *"As the situation is now, do you support peace talks between President Abbas and Israel?"*

The first question is used as the main dependent variable throughout the analysis, while the last two questions are used for robustness checks. Clearly, question (i) reflects the political preferences in the immediate aftermath of the war. The second question reflects support for the use of *violence* by Palestinians as a means to the continuation of the struggle. The third question captures the willingness of individuals to support peaceful means in resolving the conflict.

Electoral Preferences (i)

Table 2.1 shows the list of selected political parties specified for question (i). The parties are grouped together to generate an index, which is composed of two aggregate groups, moderate (M) and extreme (E). The group indexing of the parties into their respective categories are based on the party's position on the use of violence by Palestinians against Israel. Parties which are indexed with extreme supported the continuation of military actions against Israel at the time of the survey and/or have refused peaceful negotiations as a means to resolving the conflict. Given the arbitrary nature of this classification, the results may be affected by the placement of the parties along the index. Apart from Hamas and Fatah, all other listed parties only represent 2,5% of the total share of the votes. By fixing the placement of Hamas in E and Fatah in M, variations in indexing other parties did not significantly alter the main results. The real challenge, however, is indicated by the last three rows of table 2.1. About 40% of respondents either will not participate in an election if it takes place "today", refused to share information on their political preferences, or were still undecided about whom to vote for.

In order to account for this large portion of unidentifiable votes, a number methodological approaches and corrections are implemented. A multinomial logistic regression is performed enabling the pair-wise and joint analysis between all the categories of the

Table 2.1: Palestinian parties, political position, and share of votes

Party	Political Position	Group Indexing	Share of Votes Now
Fatah	Center Left	Moderate	30.6%
Hamas	Islamist (Right)	Extreme	24.6%
Mubadara	Far Left	Extreme	0.6%
Popular Front	Far Left	Extreme	1.4%
Democratic Union	Center Left	Moderate	0.4%
People's Party	Center Right	Moderate	0.1%
Other	-	-	1.4%
Not Participate	-	-	29.5%
Undecided	-	-	8.4%
Refusal	-	-	3.0 %

index, including non-participating and indecisive voters. For robustness checks, non-participation is used as a binomial dependent variable in the selection equation of a two-stage Tobit 2 model. Moreover, multiple random placement of undecided voter in both groups is undertaken for testing the validity of the results (see section 2.5 for more details on the econometric models and identification strategy).

Political Attitudes (ii) & (iii)

The Gaza War broke out because of the launching of rockets from Hamas into Israel. Table 2.2 shows the cross-tabulation frequencies between support of rocket launching and the electoral index. The first striking observation is the clear divisiveness of non-participating individuals in their support of the use of violence. Around 26% of respondents equally *strongly agree* and *strongly disagree* to such statement. The most notable trend from table 2.2 is that 49% of individuals whom will vote to an extreme party strongly disagree to the non-use of violence, while only 12% strongly agree to such actions. The reverse is also true: 44% of voters for moderate parties strongly agree to stop firing rockets, while only 16% are strongly against it. To a lesser extent, the same voting patterns apply for non-extreme cases *agree* and *disagree*, reflecting a political polarization in attitudes toward the use of violence in the Gaza Strip.

Table 2.2: Support of halting rocket launching into Israel

Question: All Palestinian factions must stop firing rockets at Israel				
	St. Agree	Agree	Disagree	St. Disagree
Moderate	235 (43.76%)	160 (29.79%)	55 (10.24%)	87 (16.20%)
Extreme	55 (11.82%)	78 (16.77%)	108 (23.23%)	224 (48.17%)
Undecided	30 (21.74%)	33 (23.91%)	30 (21.74%)	45 (32.61%)
Not Participate	132 (26.40%)	144 (28.80%)	92 (18.40%)	132 (26.40%)
All	494 (27.71%)	459 (25.74%)	314 (17.61%)	516 (28.94%)
Pearson's Chi-squared simulated p-value < 0.001				

This polarization is also evident for question (iii) on support of peace talks between Abbas and Israel as can been seen in table 2.3. 77% of moderate voters support the peace talks, while only 23% of extreme party supporters do. Moreover, there are no notable differences for the undecided group and non-participants. In total however, more individuals were likely to support peace talks between Abbas and Israel in the immediate aftermath of the war.

Table 2.3: Support of peace talks with Israel

Question: Do you support peace talks between President Abbas and Israel?		
	Yes	No
Moderate	416 (77.18%)	123 (22.82%)
Extreme	106 (23.14%)	352 (76.86%)
Undecided	69 (52.27%)	63 (47.73%)
Not Participate	275 (56.24%)	214 (43.76%)
All	953 (53.84%)	817 (46.16%)
Pearson's Chi-squared simulated p-value < 0.001		

2.4.2 Indices of Violence Exposure

Capturing the exposure of individuals to violence remains a challenge in the conflict and social sciences (Verwimp et al., 2009). Here two indicators of violence exposure are used: (i) Direct exposure, measured through the destruction of the household's own dwelling, and (ii) indirect exposure, measured through the intensity of destruction in the respondent's immediate neighbourhood.

Destruction of own dwelling: Data on own dwelling damage is solely based on reported information from the households. The question *"What is the situation of your current dwelling?"* is asked to all respondents residing in the Gaza Strip. While, the question *"What is the situation of your usual/previous dwelling?"* is only asked to households that left their homes during the war and did not return to them. There are five answer choices for both questions: {*a. completely destroyed; b. partly destroyed, cannot be rebuilt; c. partly destroyed, can be rebuilt; d. only Minor damage; e. no damage*} Based on these choices, I identify the group of households which is directly impacted by the indiscriminate bombing, and a control group which is not. Given the categorical nature of the variable, three exposure levels are generated and are used interchangeably throughout the analysis: *"Major Damage", "Minor Damage", and "All Damage (joint)"*. First, households that have answered completely destroyed or partially destroyed beyond repair are grouped together to form the new variable *Major Dwelling Damage*. Given that most households that were exposed to major dwelling damage moved their place of residence, I substitute the condition of their current dwelling with the one they lived in before the war. Second, *Minor Dwelling Damage* is constructed using households that reported repairable partial damage and minor damages to their dwellings. Lastly, *Dwell Damage All* aggregates both categories. 2.14% of households fall within the *Dwell Damage Major* category, while 48.66% fall within the *Dwell Damage Minor*, bringing the total number of households exposed to any kind of damage to 50.80%.

Table 2.4 depicts the differences for a number of household and individual characteristics between the exposed and control groups for *"Major Damage"* and *"All Damage"*. There are no notable significant differences between the two groups when using major damage as the exposure level (left-hand side of the table). Only the location of the households is different at 10% significance level, which is mainly driven by households living the Gaza governorate. This comes of no surprise given that most of the bombardment occurred in the Gaza governorate, raising the exposure level of high damage. This finding is reconfirmed in Figure 2.1 through the external satellite imagery.

The left-hand side of table 2.4 uses all (any) dwelling damages as the violence exposure level. At this level, there are a number of notable differences. First, as with *"Major Damage"*, the place of residence is significantly different between the exposed and the control group. This is to be expected given that the Northern part of the Gaza Strip (North Gaza and Gaza City) has witnessed the most intense bombardment during the war. Minor damages occurred mostly in the North Gaza governorate (also can be seen in figure 2.1). Second, the level of education is significantly different at the 5% level between the two groups. Individuals with secondary and higher education levels reported on average "no damage" than individuals with basic or no education. Third, larger households with higher number of children reported more minor damage. These three variations could reflect the targeting of dense and socio-economically weak parts of the Gaza Strip. Yet given the random nature of indiscriminate violence, and the insignificant differences in income levels, it is safe to assume that such variations are arbitrary.

Most importantly, there are no significant differences between pre-war moderate or extreme party preferences for either exposure levels. This indicates that supporters of Fatah have been equally exposed to violence (both major and minor) as the supporters of Hamas. This is likely to mitigate endogeneity concerns arising from testing the impact of violence on current electoral preferences (more details on endogeneity are discussed in section 2.5).

Destruction of Neighbourhood: The level of destruction within the neighbourhood captures the indirect exposure to violence. This variable is also constructed from the

Table 2.4: Differences in individual and household characteristics between direct exposed and control groups

	Dwelling Damage All			Dwelling Damage Major		
	No	Yes	p-value	No	Yes	p-value
N	989	1021		1967	43	
Age	35.03 (14.27)	35.14 (14.09)	0.869	35.09 (14.19)	35.02 (13.72)	0.977
Gender (Female)	49.9 (493)	47.3 (483)	0.264	48.7 (958)	41.9 (18)	0.461
Governorate % (freq)			<0.001			0.091
North Gaza	13.1 (130)	24.4 (249)		19.1 (375)	9.3 (4)	
Gaza	34.2 (338)	35.6 (363)		34.5 (679)	51.2 (22)	
Deir El-Balah	18.1 (179)	13.1 (134)		15.7 (309)	9.3 (4)	
Khan Yunis	24.8 (245)	14.4 (147)		19.6 (386)	14.0 (6)	
Rafah	9.8 (97)	12.5 (128)		11.1 (218)	16.3 (7)	
Voted E 2006	0.45 (0.50)	0.41 (0.49)	0.167	0.44 (0.50)	0.36 (0.49)	0.503
Participated 2006	0.66 (0.47)	0.67 (0.47)	0.789	0.67 (0.47)	0.58 (0.50)	0.262
Income			0.152			0.260
No income	65.2 (592)	69.2 (639)		67.3 (1207)	63.2 (24)	
< 1200 NIS	14.9 (135)	14.7 (136)		14.6 (262)	23.7 (9)	
1200-2000 NIS	11.0 (100)	9.6 (89)		10.3 (184)	13.2 (5)	
2000-3000 NIS	7.3 (66)	4.8 (44)		6.1 (110)	0.0 (0)	
> 3000 NIS	1.7 (15)	1.6 (15)		1.7 (30)	0.0 (0)	
Education			0.011			0.931
No Education	11.3 (111)	12.8 (131)		12.0 (236)	14.0 (6)	
Elementary	11.1 (109)	13.0 (133)		12.1 (237)	11.6 (5)	
Basic	22.2 (219)	25.5 (260)		23.8 (467)	27.9 (12)	
Secondary	33.3 (328)	32.0 (326)		32.7 (642)	27.9 (12)	
Tertiary	22.2 (219)	16.7 (170)		19.4 (381)	18.6 (8)	
HH Size	6.33 (2.98)	6.81 (3.19)	0.001	6.56 (3.11)	6.93 (2.69)	0.444
Employed (No)	72.6 (717)	75.5 (771)	0.157	74.1 (1457)	72.1 (31)	0.898

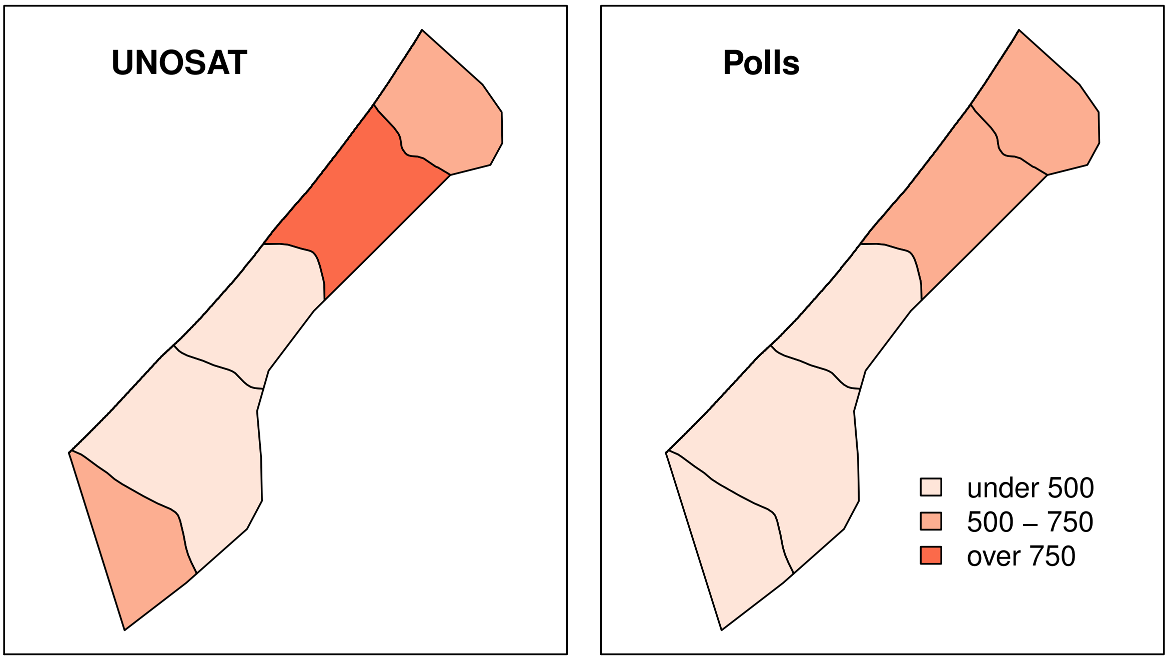
Frequencies in parentheses. The exact fisher test is used instead of the Chi squared test to account for the low number of observations within some category of the following variables: governorate, income, and education.

polls dataset and is based on the following set of questions: (i) "Has any residential house in your neighbourhood been damaged during the Gaza War? (ii) $\cdots$ any schools $\cdots$, (iii) $\cdots$ any health facility $\cdots$, and finally (iv) $\cdots$ any mosque $\cdots$." The possible three answer choices vary from "completely damaged" to "no damage at all". As for the last three facilities, a fourth answer choice is added which states: "there is no such facility in my neighbourhood."

A binary variable "Neighbourhood Damage All" takes the value of 1 if at least one of the facilities in the neighbourhood listed above is reported to have any kind of damage, and 0 otherwise. On the other hand, the binary variable "Neighbourhood Damage Major" only includes respondents who reported at least having a fully destroyed facility, or one that is damaged beyond repair, in their area of residence. The binary transformation of the variable is used to produce comparable results to the "Own Dwelling Damage" variable. The downside of such transformation is that it disregards the intensity of violence in the neighbourhood. In other words, destruction of all facilities carries the same weight as a destruction of one facility. Alternatively, one can use of simple sum of all damaged facilities to better capture the variation in intensity. However, such a simple method disregards the absence of a facility in the area of the respondent, and is endogenous to the location of the household (more dense areas are likely to have more facilities). Therefore, in order to reduce any error in the estimates by using an aggregate damage index, and produce comparable coefficients, the binary transformation seems to be the most appropriate choice.

The satellite data from UNOSAT is used to check the reliability of the respondents' answers. Given that the polls data does not contain the exact coordinates of the households, the comparison is only undertaken at the governorate level. Figure 2.2 shows the intensity of violence at the governorate level from both sources. The left-hand side of figure 2.2 depicts the level of building destruction captured by the satellite imagery of UNOSAT, while the right-hand side of the figure shows the estimated frequency of

Figure 2.2: Location damage in Gaza - governorate level



Note: *Only destroyed building are used for plotting the geo points of the UNOSAT sub-figure.*

reported major neighbourhood destruction from the polls dataset.[5] There are close similarities between the two spatial distributions, with a slight error measurement in the most southern governorate. The estimated frequency of the destruction of the polls data does not account for the dispersion of the population. Hence there is a possibility of over- or under-counting the actual incidences of destruction.[6] In other words, households in dense areas living in proximity to one another could have reported the same destruction points in the UNOSAT data.

Table A.1 in appendix of chapter 2 shows the differences of several household individual characteristics between the exposed and control group for both major and any neighbourhood damage. Similar to the results for variable of direct exposure to violence, the location variable is significantly different for both exposure levels. Respondents who

5 *The estimated frequency is calculated on a three-step basis. First, extracting the percentage of respondents reporting major destruction. Second, extracting the distribution of those respondents by governorate. Third, extrapolating the percentage distribution to the total number of data points from UNOSAT.*

6 *This is one weakness of cluster sampling which disregards whole enumeration areas from the sample.*

reported any damage in their immediate neighbourhood are more likely to be supporters of moderate parties. This reconfirms the assumption that violence was indiscriminate and did not specifically target Hamas supporters. Yet, the 5% level of significance might impose bias in the analysis and the estimates should be interpreted carefully using this indirect exposure level.

2.5 Empirical Framework

2.5.1 Identification Strategy

Two different control groups are selected for the analysis. First, the comparison is conducted between households with damaged dwellings, both fully and partially, as well as non-damaged dwellings within the Gaza Strip. Estimating the impact of destruction on the political preferences and voting outcomes remains difficult to disentangle due to the possible self-selection of areas of bombardment. This is reflected by the concentration of the bombardment to the northern and border areas of the Gaza Strip, as well as around Gaza City as demonstrated in figure 2.1. The self-selection is problematic particularly if households in the the affected areas are initially supportive to the extremist group Hamas. In order to overcome this issue of selection of the treatment and control group within Gaza Strip, the counterfactual should predict random selection of affected households. That is, the causal estimates will only be unbiased when the bombing (destruction of dwellings) is considered exogenous regarding the political preferences of households.

From a military point of view, the precision of the air strikes of the Israeli army is considerably accurate. Therefore targeted bombing is more likely to be used during this war. Nevertheless, the nature of the unequal military arsenal prompted Hamas fighters to use locations within residential neighbourhoods for launching their rockets towards Israel. Given that the retaliation of the Israeli air force is quick and precise against these specific launching locations, one can then assume that households within these residential areas are unaware to the rocket launching taking place. Therefore, from a pure tactical point of view, the targeted rockets on residential areas are considered random shocks to households. On the other hand, one can counteract such argument by proposing that households that support Hamas are more prone to accept rocket launching in the vicinity of their residential areas. Although the argument does sound extreme, given that children, women, and elderly, were killed during these bombings, one can easily test statistically for such an assumption. As shown in table 2.4, there are no significant differences of exposure to bombing between households supporting Hamas or Fatah before the start of the war.

2.5.2 Econometric Model

Given the nature of the main dependent variable, there are analytically two feasible ways forward. The first option is to break up the categorical dependent variable into binaries and run multiple logistic regressions to estimate the impact of violence on voting preferences. The second approach retains the categorical nature of dependent variable where a multinomial logit regressions is performed to tease out the differences between the various political preferences.

Binomial Model - Dummy Approach

First, I estimate the likelihood of voting for extreme parties by compacting the categorical dependent variable into several binomial variables as follows:

$$\bar{Y}_{ij} = \begin{cases} 1 & \text{if } Y_{ij} = \text{j} \quad \forall j \in J \\ 0 & \text{otherwise} \end{cases}$$

where i resembles the ith individual J is the set of factor levels of Y, such that J : $\{E, M, O, NP, U\}$[7]. By generating several dummy variables it is easier to explicitly measure the impact of violence on singular voting preferences while maintaining the original number of observations. Although such approach generally increases the model's

7 *E=Extreme; M=Moderate; NP=Not Participating; U=Undecided*

goodness of fit, grouping all other levels together does not permit estimating the changes within the voting index. Hence, using $\bar{Y}$ as the main dependent variable, I am able to examine the impact of violence on voting extreme versus all other options by applying a maximum likelihood estimation of the following logit model:

$$\bar{Y}_{it} = \alpha_0 + \alpha_1 \cdot Dest_{it} + \alpha_2 \cdot \bar{Y}_{i\bar{t}} + \alpha_2 \cdot \bar{Y}_{i\bar{t}} \times Dest_{it} + \alpha_4 \cdot Z_{it} + \epsilon_{it} \tag{2.1}$$

where $Dest_{it}$ takes one of the destruction indices as developed in section 2.4, and $Y_{i\bar{t}}$ captures the actual election choice of individual i in the 2006 Gaza election.[8]. The coefficient α_2 captures the impact of destruction of individuals who had similar party preferences in the 2006 election. Analysis with and without the the interaction term are presented.

Multinomial Model

The second option retains the categorical nature of the dependent variable with the four distinct categories as presented in J. Let

$$\pi_{ij} = Pr\{Y_i = j\}$$

denote the probability that the i-th response falls within the j-th category. A multinomial logistic regression is performed to tease out the differences between the likelihood of individuals choosing between the various factor levels. The use of multinomial logit models is widespread in the analysis of voting and polls data given that categorical variables depicting voting preferences are mutually exclusive and exhaustive - which are two vital criteria to produce consistent estimates. The log-odds response of each category

[8] *with similar levels of J excluding the levels U*

follows a linear model:

$$Election_{ijt} = log(\frac{\pi_{ij}}{\pi_{iE}}) = \beta_{0j} + \beta_{1j}Dest_{it} + \beta_{2j}Election_{i\bar{j}\bar{t}} + \beta_{3j}Election_{i\bar{j}\bar{t}} \times Dest_{it} + \alpha_{4j}Z_{it} + \epsilon_{ijt} \tag{2.2}$$

where i is the i-th individual and j is the j-th category of the dependent variable. Extreme is used as the baseline category with $j_{baseline} = E$. Moreover the actual election preferences in 2006 are depicted $Election_{i\bar{j}\bar{t}}$ where $\bar{j} \in \bar{J} : \{E, M, NP, O\}$, and $\bar{t} = 2006$. Most importantly the β_{nj} is a vector of regressions coefficients, for $j = 1, 2, ..., J-1$. The model is closely similar and analogous to the logistic regressions model of equation 2.1, but the probability distribution of the response is multinomial instead of binomial with $J-1$ equations instead of one. In other words, equation 2.2 contrasts each category with $j_{baseline}$ while equation 2.1 contrasts only extreme and non-extreme voters.

2.6 From Violence to Voting

The first part of the analysis examines how exposure to both direct and indirect violence affects the voting choices of individuals. The results are based on the empirical estimations of both binomial and multinomial logistic regressions. Moreover, robustness checks on the inclusion of non-participation in election in the outcome variable is performed.

2.6.1 Voting Extreme

Table 2.5 shows the impact of violence exposure on the likelihood of voting to extreme parties. The estimations include the core variables of interest (damage and previous election preferences), and a vector of explanatory individual and household characteristics. Individual characteristics constitute age, gender, educational level, and socio-economic status (SES) of respondents. The SES is measured via the variables of current employment status and the income levels in the past month. Moreover, an index of psychological status is also included to assess the emotional well-being of respondents. The index is developed by aggregating ten questions capturing anxiety, restlessness, depression, hopelessness, and loss of appetite. Household size and number of children capture the main household characteristics. Most importantly, I control for households that have received any income from Hamas one week before the interview. The variable is envisioned to influence party choices of respondents.

Direct Exposure: Columns (1) to (4) of table 2.5 present the results for own dwelling damage. Columns (1) and (2) use all (any) reported damage by the respondents, while (3) and (4) use only major damage. Moreover, all models are presented with and without controlling for the interaction term between damage and prior election preferences.

First, direct exposure to any kind of destruction has a significant negative effect on voting for extreme parties. As visible in column (2), the impact remains significant at the 1% level even after the inclusion of the interaction term. However, the interaction term is not significant itself and has a positive sign. This implies that the impact of destruction does not convert prior extreme party supporters towards other preferences.

Second, individuals exposed to direct high intensity of violence, measured through major damage to own dwelling, are also more likely not to have extreme electoral preferences. The effect is significant at the 10% level, yet it disappears when the interaction terms is controlled for as shown in column (4).

Moreover, females and younger people are more likely to vote for extreme parties. Respondents who voted for an extreme party in the previous elections and respondents who have received monetary assistance from Hamas after the war are significantly likely to continue voting for extreme parties regardless of the exposure levels of violence. There is no effect of education and SES variables in increasing the likelihood to vote for extreme political parties.[9]

Indirect Exposure: There are no effects of neighbourhood damage (both for major and any type of damage) on the political preferences of individuals, as seen in columns (5)

[9] *Income and Education levels are repressed in the table to save space.*

Table 2.5: Impact of direct and indirect violence exposure on voting for extreme parties

	Own Dwelling Damage				Neighborhood Dwelling Damage			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Vote E 2006	1.46***	1.37***	1.45***	1.45***	1.44***	1.42***	1.44***	1.47***
	(0.08)	(0.11)	(0.08)	(0.08)	(0.08)	(0.13)	(0.08)	(0.09)
Damage All	−0.25***	−0.34***			−0.11	−0.12		
	(0.08)	(0.10)			(0.08)	(0.10)		
Vote E 2006*Damage All		0.20				0.03		
		(0.15)				(0.16)		
Damage Major			−0.59*	−0.48			−0.01	0.04
			(0.31)	(0.39)			(0.08)	(0.11)
Vote E 2006*Damage Major				−0.27				−0.10
				(0.62)				(0.17)
Age	−0.01***	−0.01***	−0.01***	−0.01***	−0.01***	−0.01***	−0.01***	−0.01***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Gender (Female)	0.18**	0.19**	0.19**	0.19**	0.21**	0.21**	0.21**	0.21**
	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)
Employed (No)	0.15	0.15	0.14	0.14	0.12	0.12	0.13	0.13
	(0.14)	(0.14)	(0.14)	(0.14)	(0.14)	(0.14)	(0.14)	(0.14)
Income Hamas	0.77***	0.77***	0.79***	0.79***	0.79***	0.79***	0.79***	0.79***
	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)
Emotional Instability	−0.01**	−0.01**	−0.01***	−0.01***	−0.01***	−0.01***	−0.02***	−0.02***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
HH Size	0.05***	0.04***	0.04**	0.04**	0.04**	0.04**	0.04**	0.04**
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
N child	−0.05**	−0.05**	−0.05**	−0.05**	−0.05**	−0.05**	−0.05**	−0.05**
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
BIC	1629.90	1635.60	1637.25	1644.49	1630.20	1637.61	1632.10	1639.18
Log Likelihood	−748.00	−747.13	−751.67	−751.58	−748.19	−748.18	−749.14	−748.96
Num. obs.	1701	1701	1701	1701	1693	1693	1693	1693

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard Errors in Parenthesis. voting Moderate is used as the baseline category in the multinomial regression analysis for both the depedenet variable and the Election variable of 2006. All other sub-categories (Not Participate, Others, Undecided) are not included in the table. Moreover, only significant income and education categories are shown. Employment, Number of children, and the constant term are repressed.

to (8). The main coefficients inhibit similar signs to the ones of direct exposure to violence, except for specification (8), where dwellings in the neighbourhood are completely damaged.

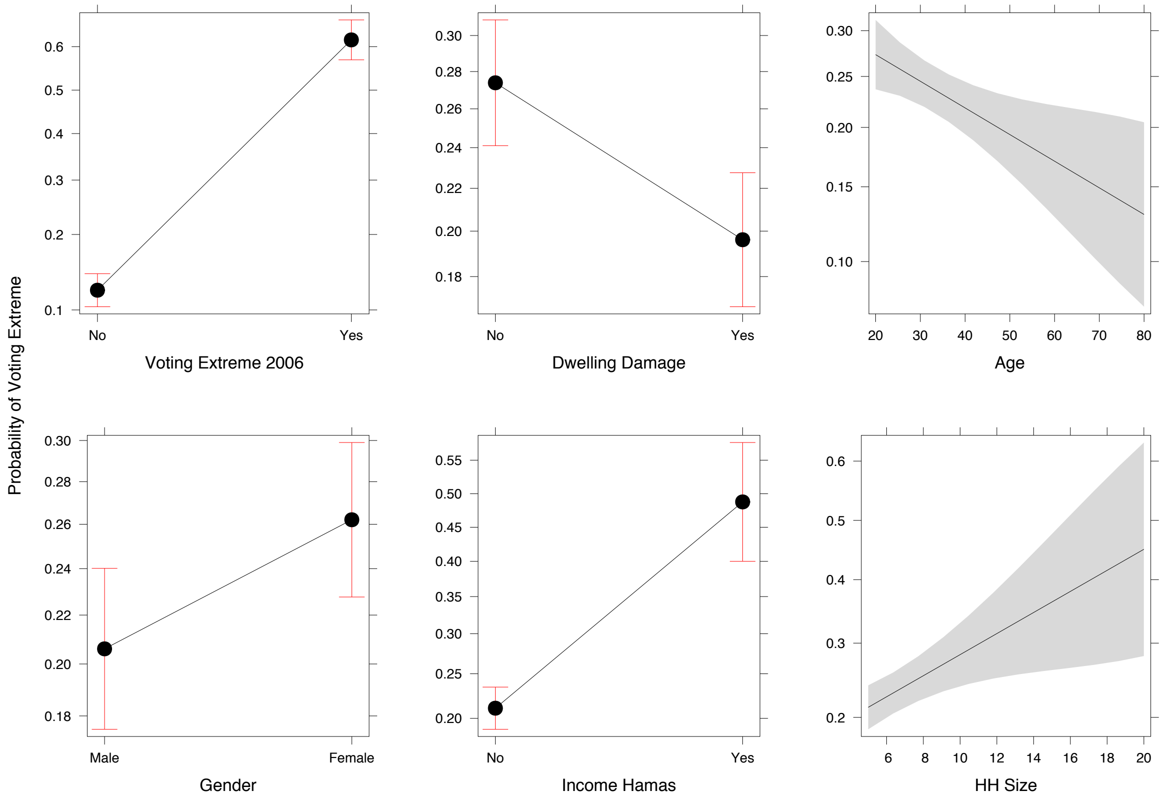
The coefficients are not directly self-explanatory given the nature of the maximum likelihood estimations used. Hence, it is preferable to interpret the results through the marginal effects. In order to better understand the impact of the independent variables on the likelihood of voting for extreme parties, figure 2.3 presents the marginal effects on the probability of voting extreme. For better interpretation of the coefficients of the main regression, the figures show the average adjusted prediction at representative values for the main significant variables in the model. The predictive effects are presented for each variable separately while holding all other variables at their observed values.

It is clear that prior voting preferences plays the most crucial role in determining the voting choices today. Individuals who did not vote for extreme parties in 2006 have a very low likelihood of voting to extreme parties at the time of the interview. More than 60% of respondents will vote extreme if they had the same party preferences in 2006. Moreover, receiving income from Hamas directly after the end of the war strongly determines the probability of voting for extreme parties. Surprisingly, females, younger people, and large households are more likely to vote for extreme parties, with probabilities of 26%, 28%, and 44%, respectively.

Being exposed to violence, measured via any dwelling damage inflicted on the household, reduces the likelihood of voting to extreme parties, although the difference in impact is not strongly significant. Individuals from households that were exposed to damage exhibit a 20% probability of having radical political preferences, while the figures for households that were not exposed to violence is slightly larger at 27%.

Figure 2.4 compares the marginal effects of the main explanatory variable (damage) from the four specifications of table 2.5 which exclude the interaction term. The y-axis is benchmarked on the response probability of voting extreme to make the sub-figures

Figure 2.3: Marignal effects on the probability of voting extreme

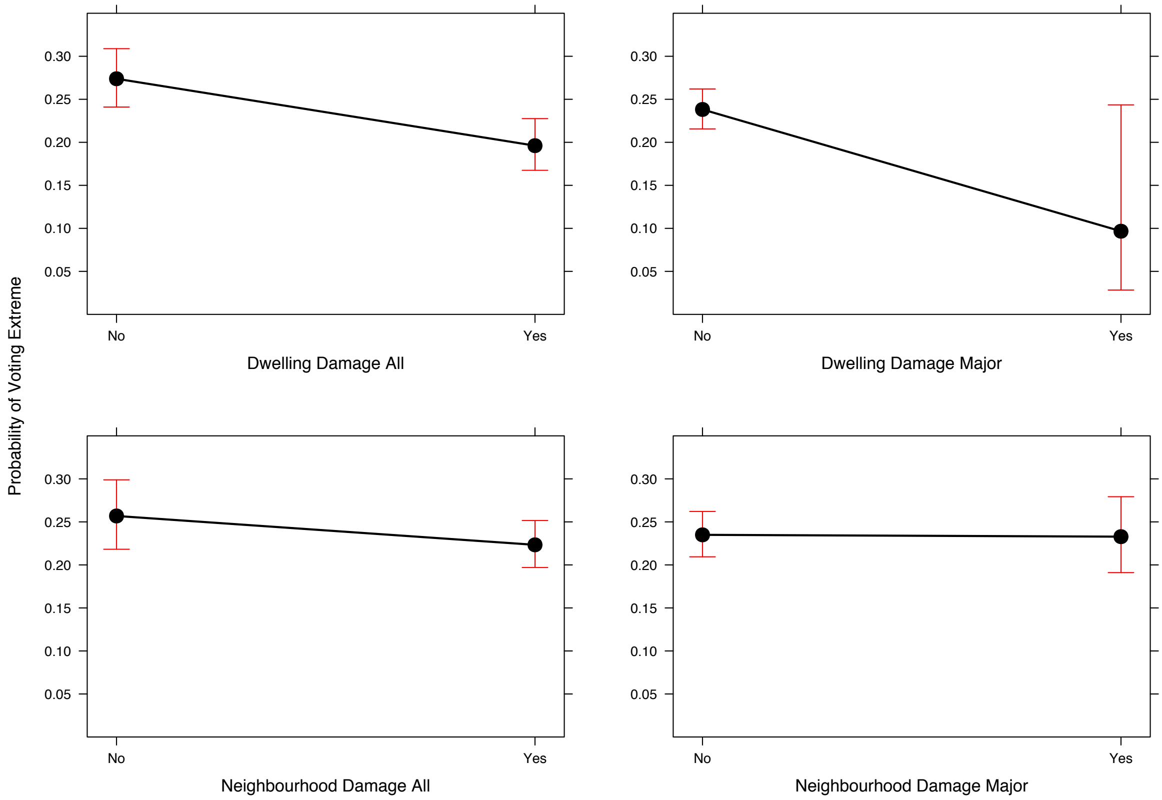


Note: *Only variables with significance levels above 10% from column (1) table 2.5 are shown. Dummy variables are transformed into factors for clarity of presentation.*

visibly comparable. As before, although the effect of any direct exposure to violence is less pronounced than major direct exposure, the significance of the estimators are stronger. This can be seen via the large confidence intervals for the latter group, which is reflected through the 10% significance levels. Moreover, the number of observations are much smaller compared to the first group, hence the actual effect might be overestimated. On the other hand it is clear from the lower section of the figure 2.4 that indirect exposure to violence has no impact on the voting preferences.

The results suggest that exposure to direct violence repels individuals from supporting radical groups. Yet the impact is only significant for individuals that experienced any kind of damage, which disregards the intensity of the exposure. On the other hand, individuals that had their homes destroyed or damaged beyond repair are less prone to

Figure 2.4: Various effects of exspoure on the probability of voting extreme



Note: *Damage variable from columns (1) , (3), (5), (7) respectively of table 2.5 are shown. Dummy variables are transformed into factors for clarity of presentation.*

support extreme parties, yet the effect disappears when previous electoral preferences are accounted for. In other words, exposure to violence has a polarization effect on electoral preferences, alienating non-extreme individuals further away from extreme groups, while having no notable impact on prior extreme party supporters. These outcomes contradict the results of Jaeger et al. (2012), where they find that indirect exposure to violence initially attracts individuals towards radicalisation, but where the effect diminishes with time. Unfortunately, there is no possibility to test the long-term impacts of violence on individuals in the current dataset. Nevertheless, it would be interesting to check if the divergence of individuals from extreme parties is sustained for longer periods of time.

2.6.2 Voting Moderate

Next, similar estimations of equation 2.1 are used to determine the impact of violence on voting to moderate parties. Again, given the binary transformation of the dependent variable, it is difficult to clearly understand changes in voting patterns. For example, voters that deterred from voting for extreme parties do not necessarily shift their preferences towards moderate parties. This is true given that the zeros of the dependent variable in the first estimation also included the categories *"Not Participate"*, and *"Undecided"*.

Table 2.6 presents the results for voting moderate with the same model specifications as in the initial estimations. Exposure to any kind of damage both directly and indirectly has an impact on individuals in voting for moderate parties. Moreover, individuals with tertiary education (benchmarked to individuals with no education) are less likely to vote to moderate parties. Although the channels behind such an effect are hard to disentangle with this level of analysis, previous research found similar effects (Krueger and Malečková, 2003). More educated individuals understand the complexity of the Palestinian-Israeli conflict better, reiterating the long-standing failure of Palestinian moderate parties in resolving the conflict. Given that the affect was not evident for voting to extreme parties, it is safe to assume that individuals with tertiary education are becoming more apolitical. This reasoning can also be applied to understand the negative impact of age on neither voting to extreme nor moderate parties compared to the younger cohorts. Contrary to the results of table 2.5, individuals with high income levels (> 2000 NIS) are more likely to vote to moderate parties compared to individuals reporting no income.

The impact of violence exposure extends to the indirect damage of the neighbourhood. Individuals who witnessed any kind of damage in their area are more likely to support moderate parties. However, the effect is not strongly significant and does not extend to areas of high intense damage. To that end, the use of logistic regression on both

Table 2.6: Impact of direct and indirect violence exposure on voting for moderate parties

	Own Dwelling Damage				Neighborhood Dwelling Damage			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Vote M 2006	1.68***	1.77***	1.67***	1.66***	1.67***	1.73***	1.67***	1.69***
	(0.08)	(0.11)	(0.08)	(0.08)	(0.08)	(0.14)	(0.08)	(0.09)
Damage All	0.28***	0.35***			0.16**	0.20*		
	(0.08)	(0.10)			(0.08)	(0.11)		
Vote M 2006*Damage All		−0.18				−0.09		
		(0.15)				(0.16)		
Damage Major			0.50**	0.38			−0.02	0.01
			(0.25)	(0.30)			(0.09)	(0.11)
Vote M 2006*Damage Major				0.51				−0.09
				(0.63)				(0.17)
Age	−0.02***	−0.02***	−0.02***	−0.02***	−0.02***	−0.02***	−0.02***	−0.02***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Gender (Female)	−0.10	−0.10	−0.10	−0.11	−0.11	−0.11	−0.11	−0.11
	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)
Education (Tertiary)	−0.24	−0.25	−0.26*	−0.27*	−0.27*	−0.27*	−0.28*	−0.28*
	(0.16)	(0.16)	(0.16)	(0.16)	(0.16)	(0.16)	(0.16)	(0.16)
Income (2000-3000 NIS)	0.45**	0.46**	0.44**	0.45**	0.48***	0.48***	0.46**	0.45**
	(0.18)	(0.19)	(0.18)	(0.18)	(0.19)	(0.19)	(0.18)	(0.18)
Income (> 3000 NIS)	0.52*	0.52*	0.54*	0.54*	0.54*	0.54*	0.53*	0.54*
	(0.29)	(0.29)	(0.29)	(0.29)	(0.29)	(0.29)	(0.29)	(0.29)
Income Hamas	−0.65***	−0.65***	−0.67***	−0.67***	−0.65***	−0.64***	−0.65***	−0.65***
	(0.17)	(0.17)	(0.17)	(0.17)	(0.16)	(0.16)	(0.16)	(0.16)
Emotional Instability	0.00	0.00	0.01	0.01	0.01	0.01	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
BIC	1599.07	1605.08	1608.38	1615.10	1604.72	1611.87	1608.80	1615.99
Log Likelihood	−732.58	−731.87	−737.24	−736.88	−735.45	−735.31	−737.49	−737.37
Num. obs.	1701	1701	1701	1701	1693	1693	1693	1693

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard Errors in Parenthesis. voting Moderate is used as the baseline category in the multinomial regression analysis for both the dependent variable and the Election variable of 2006. All other sub-categories (Not Participate, Others, Undecided) are not included in the table. Moreover, only significant income and education categories are shown. Employment, Number of children, and the constant term are repressed.

moderate and extreme binary variables seem to suggest that their is a willingness for voters to support moderate parties over extreme ones if they were exposed to violence. However, these results do not project the full picture clearly, giving the difficulty in detecting the interplay between the various electoral sub-categories.

2.6.3 Results from the Multinomial Approach

The multinomial regression estimates provide a clearer picture on the impact of violence on voting to extreme parties *vis-á-vis* moderate ones. Table 2.7 shows the results of the multinomial estimates on voting to extreme parties with moderate as the baseline category. All other sub-categories are disregarded in the table for a better clarity in the presentation of results. As seen in columns (1) to (4), the coefficients of damage (both major and all) are significant with a negative sign. This suggests that voters are more likely to support moderate parties over extreme ones if they have been exposed to destruction during the war. This finding reconfirms the results from the initial logit estimations. Interestingly, the interaction term between major damage and voting extreme over moderate in 2006 is positive and significant at the 1% level - column (4).

This result implies that despite the average effect of de-radicalization of exposure to intense direct violence, voters who already support extreme parties will continue doing so. This reconfirms the polarization effect captured in the first part of the analysis, and highlighted by Jaeger et al. (2012). However, in contrast to Jaeger et al. (2012), this polarization effect is driven initially via a de-radiclization impact. The goodness of fit estimates suggest that the current specification performs worse than the logit model. This is to be expected given the small number of observations within each category of the dependent variable. Yet, given that the model has converged properly and that it has not violated any of the initial criteria of multinomial logit estimations, it is safe to assume that the estimates are unbiased.

The results for voting for moderate parties is shown in table A.2 in the appendix of this chapter. The effect, as expected, is reversed. The coefficients of the columns without the

Table 2.7: Impact of direct and indirect violence on voting extreme- multinomial appraoch

	Own Dwelling Damage				Neighborhood Dwelling Damage			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Vote E 2006	5.39***	4.95***	5.32***	5.36***	5.32***	5.19***	5.31***	5.47***
	(0.32)	(0.41)	(0.31)	(0.32)	(0.31)	(0.52)	(0.31)	(0.37)
Damage All	−0.75***	−1.31**			−0.38**	−0.33		
	(0.17)	(0.53)			(0.18)	(0.47)		
Vote E 2006*Damage All		1.07				0.18		
		(0.68)				(0.65)		
Damage Major			**−1.56****	**−11.31*****			0.00	0.29
			(0.65)	**(0.46)**			(0.19)	(0.51)
Vote E 2006*Damage Major				**8.77*****				−0.58
				(0.75)				(0.68)
Age	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Gender (Female)	0.42**	0.44**	0.42**	0.42**	0.47**	0.47**	0.45**	0.45**
	(0.20)	(0.20)	(0.20)	(0.20)	(0.20)	(0.20)	(0.20)	(0.20)
Education (Basic)	0.61*	0.65*	0.56	0.53	0.61*	0.61*	0.60*	0.60*
	(0.35)	(0.35)	(0.35)	(0.35)	(0.35)	(0.35)	(0.35)	(0.35)
Education (Secondary)	0.58*	0.62*	0.57*	0.55	0.61*	0.61*	0.59*	0.59*
	(0.35)	(0.35)	(0.35)	(0.35)	(0.35)	(0.35)	(0.35)	(0.35)
Income (2000-3000 NIS)	−0.82*	−0.82*	−0.80*	−0.80*	−0.93**	−0.91*	−0.88*	−0.87*
	(0.46)	(0.46)	(0.46)	(0.46)	(0.47)	(0.47)	(0.47)	(0.47)
Income Hamas	1.74***	1.77***	1.79***	1.82***	1.75***	1.76***	1.75***	1.75***
	(0.36)	(0.36)	(0.36)	(0.37)	(0.36)	(0.36)	(0.36)	(0.36)
Emotional Instability	−0.02	−0.02	−0.03**	−0.03**	−0.02**	−0.02**	−0.03**	−0.03**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
BIC	4319.68	4397.51	4333.55	4414.33	4315.82	4396.37	4321.75	4399.98
Log Likelihood	−1862.28	−1856.56	−1869.22	−1864.97	−1860.54	−1856.21	−1863.51	−1858.01
Num. obs.	1701	1701	1701	1701	1693	1693	1693	1693

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$. Standard Errors in Parenthesis. voting Moderate is used as the baseline category in the multinomial regression analysis for both the depedenet variable and the Election variable of 2006. All other sub-categories (Not Participate, Others, Undecided) are not included in the table. Moreover, only significant income and education categories are shown. Employment, Number of children, and the constant term are repressed.

interaction term are the exact opposites of table 2.7 (columns (1), (3), (5), and (7)). The results of the interaction terms as presented in bold in column (4) show the significance of the polarization effect in driving the results behind the estimations. Both moderate and extreme party supporters are more likely to continue supporting their political factions, respectively, regardless of the magnitude of violence exposure. Yet, even after controlling for the interaction term, there is an immediate de-radicalization trend among individuals exposed to violence.

2.7 Robustness Checks

2.7.1 Sample Selection Bias

Electoral party preferences are observed if and only if the respondents take part in an election. Therefore, individuals deciding not to partake in the election are likely to constitute a non-random sub-sample of all respondents eligible for the RSI questionnaire. Throughout the main analysis, voters and non-voters have been treated similarly. Both binomial and multinomial estimates included non-voters in the outcome variables. This could pose a bias given that the determinants of voting might differ from that of participation. In order to overcome such a challenge, and to test the robustness of the results, two approaches are implemented. First, the analysis is undertaken only on the sub-sample of individuals participating in the elections, ignoring concerns over non-random selection mechanisms. In other words, all individuals that chose not to vote are excluded from the analysis. The advantage in cropping the sample lies in providing more accurate estimates on voting preferences. The main disadvantage of such a method is the lower number of observation, which nullifies the representativeness of the sample.

The second approach accounts for sample selection bias of voters. If the drivers of abstention are significantly different from that of choosing a moderate or an extreme party, then the estimates are likely to be inconsistent. In order to account for a possible selection bias, I estimate a two-stage Tobit II model with sample selection correction. The model takes into account the probable correlation of omitted unobservables impacting both election participation and electoral choices.

In the first stage, I estimate the selection equation by running a probit regression on the determinants of participation as in equation 2.1, such that $j = 1 - NP$

$$Partic_{it} = \alpha_0 + \alpha_1 \cdot Partic_{i2006} + \alpha_2 \cdot X_{it} + \epsilon_{it} \tag{2.3}$$

$Partic_{it}$ is a binomial variable taking the value 1 if individuals are willing to participate in an election, and 0 otherwise. $Partic_{i2006}$ is the actual participation rates in the 2006 election. X_{it} are relevant household and individual characteristics. Apart from age and gender, socio-economic status has been widely demonstrated in literature to play an important role on election abstention, which includes level of education, income, and employment status (e.g., Bühlmann and Freitag, 2006). Moreover, recent evidence shows that psychologically healthy individuals are more likely to be politically active and to vote(Gerber et al., 2011). A Likert scale of life satisfaction variables is included only in the first stage as a determinant of participation to account for the exclusion restriction criterion of the model. As in previous estimations, an aggregate proxy reflecting the current mental health of respondents is included as well.

In the second stage, I then estimate the impact of destruction on voting for extreme parties as set in equation 2.4. The regression includes party preferences in the 2006 election, as well as the interaction term between destruction and the 2006 preferences. λ is the Inverse Mill's Ratio (IMR) obtained from the first stage equation. Given the binomial nature of the outcome variable, it is inefficient to use OLS estimation in the two-step process. Instead the ML estimates based on the Newton-Raphson maximization procedure are performed.

$$E_{it} = \beta_0 + \beta_1 \cdot Dest_{it} + \beta_2 \cdot E_{it-1} + \beta_2 \cdot E_{it-1} \times Dest_{it} + \beta_4 \cdot Z_{it} + \lambda + \epsilon_{it} \tag{2.4}$$

Table 2.8: Tobit II model for sample selection: impact of violence on voting for extreme parties

	Own Dwelling Damage				Neighborhood Dwelling Damage			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Vote E 2006	1.79***	1.61***	1.78***	1.79***	1.77***	1.65***	1.77***	1.80***
	(0.11)	(0.14)	(0.11)	(0.11)	(0.10)	(0.16)	(0.10)	(0.11)
Damage All	−0.30***	−0.43***			−0.08	−0.15		
	(0.09)	(0.11)			(0.09)	(0.11)		
Vote E 2006*Damage All		0.37**				0.20		
		(0.19)				(0.20)		
Damage Major			−0.73**	−0.59			0.04	0.07
			(0.35)	(0.43)			(0.10)	(0.12)
Vote E 2006*Damage Major				−0.37				−0.09
				(0.70)				(0.21)
Age	−0.01	−0.01	−0.01	−0.01	−0.01	−0.01	−0.01	−0.01
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Gender (Female)	0.35***	0.36***	0.34***	0.34***	0.37***	0.37***	0.37***	0.36***
	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)
Education (Basic)	0.35*	0.37**	0.33*	0.33*	0.35*	0.36*	0.36**	0.36**
	(0.18)	(0.18)	(0.18)	(0.18)	(0.18)	(0.18)	(0.18)	(0.18)
Income Hamas	0.98***	0.99***	0.99***	0.99***	0.97***	0.97***	0.97***	0.97***
	(0.16)	(0.16)	(0.16)	(0.16)	(0.16)	(0.16)	(0.16)	(0.16)
Emotional Instability	−0.02**	−0.02**	−0.02***	−0.02***	−0.02***	−0.02***	−0.02***	−0.02***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
HH Size	0.05**	0.05**	0.05**	0.05**	0.04*	0.05**	0.04**	0.04**
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
BIC	3300.65	3304.21	3307.48	3314.64	3290.71	3297.12	3291.35	3298.62
Log Likelihood	−1520.21	−1518.28	−1523.63	−1523.49	−1515.33	−1514.82	−1515.65	−1515.57
Censored	485	485	485	485	478	478	478	478
Observed	1209	1209	1209	1209	1208	1208	1208	1208

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$. Standard Errors in Parenthesis. Results from first stage equation are not shown. Only significant income and education categories are shown. All other insignificant covariates are repressed.

Table 2.8 shows the results after accounting for sample selection bias.[10] The estimates are based on the maximum likelihood of voting to extreme parties versus all others, including moderate and undecided voters. Only the results of the second stage outcome equation are presented in table 2.8. There are no significant variations from the original logistic estimations after censoring the sample to only those who are willing to participate in elections. Direct major damage has a negative effect on voting for extreme parties. Yet the impact disappears when controlling for extreme party preferences in 2006. The results remain significant for all direct forms of damage, highlighting the positive coefficient of the interaction term in column (2). Voters who already support extreme parties are more likely to continue supporting them if they are exposed to violence. On the other hand, there is no impact of indirect exposure to violence as seen in columns (4) to (8). To summarize, intense major damage does not have a de-radicalization effect, but rather a further polarization effect among the electorate even after controlling for sample selection bias.

2.7.2 The Undecided Group

8% of the respondents are undecided on whom to vote for if an election is to take place. Moreover 10% of the undecided group have been exposed to major direct violence and had their dwelling completely destroyed. Literature on polls analysis has been conflicted on how to deal with the undecided group. The aim of the robustness check is to perform a random and conditional assignment of the undecided category into the other remaining groups. By doing so, the likelihood that this group might play in producing unbiased results on the impact of violence on voting preferences is mitigated. Although this category is separately examined in the multinomial regression model, it remains unclear how mutually exclusive this category is from the remaining three, and hence the criterion of producing consistent and unbiased estimates is still not satisfied.

In order to reduce the likelihood that those 10% can impact the direction and signific-

[10] *Results on the simple sub-sample procedure can be found in appendix of chapter 2.*

ance of the results, I undertake a number of robustness checks. These robustness checks mainly involve reassigning the undecided group into the other three categories of J'. Yet such an assignment is not an easy task given that it can endogenously impact the results. Hence, various reassignment techniques are implemented and cross-validated. First, by replacing all the undecided voters into either extreme or moderate category (extreme bounds). Second, by randomly assigning the undecided group into the remainder categories with random probabilities of replacement. Third, by randomly assigning the group into the remainder categories with equal probabilities. Fourth, by randomly assigning the undecided group into the remainder categories with probabilities equal to the distribution of the votes from the 2006 election. Finally, by replacing the undecided group with the remainder categories such that the distribution is exactly equal to the preferences of individuals in the 2006 election.

Under all these variable reassignments, the results of the main coefficients of interest are unaffected. This is true for all levels of exposure to violence. Therefore it safe to conclude that the undecided group does not have a potent role on the direction and significance of the results.[11]

2.7.3 Political Attitudes as Dependent Variables

In order to ensure that results are not driven by the nature of the dependent variable of electoral preferences in the dataset, the estimations are rerun using various indicators of extremism. These indicators are extracted from the two other variables on political attitudes described in section 2.4.1. The first variable is based on the question asks respondents if they support the peace talks between Abbas and Israel. Logistic regressions are estimated on the Yes/No dependent variable, where the value 1 is given to individuals who answered yes. Table 2.9 shows the results of the estimation using direct and indirect measures of damage exposure. In line with results of electoral preferences, there is no significant impact of damage on moderate opinion towards the peace talks. Only,

[11] *The tables of the results of the robustness checks of the undecided group are available upon request.*

Table 2.9: Impact of violence on the support of peace talks

	Own Dwelling Damage		Neighbourhood Damage	
	(1)	(2)	(3)	(4)
Damage All	0.09		0.16^{**}	
	(0.06)		(0.07)	
Damage Major		-0.15		-0.00
		(0.22)		(0.07)
Age	-0.00	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)
Gender (Female)	0.06	0.05	0.05	0.05
	(0.07)	(0.07)	(0.07)	(0.07)
Education (Elementary)	-0.26^{*}	-0.26^{*}	-0.26^{*}	-0.27^{*}
	(0.14)	(0.14)	(0.14)	(0.14)
Education (Basic)	-0.42^{***}	-0.41^{***}	-0.41^{***}	-0.41^{***}
	(0.13)	(0.13)	(0.13)	(0.13)
Education (Secondary)	-0.53^{***}	-0.53^{***}	-0.54^{***}	-0.54^{***}
	(0.13)	(0.13)	(0.13)	(0.13)
Education (Tertiary)	-0.46^{***}	-0.46^{***}	-0.46^{***}	-0.47^{***}
	(0.13)	(0.13)	(0.13)	(0.13)
Income Hamas	-0.69^{***}	-0.68^{***}	-0.70^{***}	-0.70^{***}
	(0.11)	(0.11)	(0.11)	(0.11)
Emotional Instability	0.01^{***}	0.01^{***}	0.01^{***}	0.01^{***}
	(0.00)	(0.00)	(0.00)	(0.00)
HH Size	-0.04^{***}	-0.04***	-0.04^{***}	-0.04^{***}
	(0.01)	(0.01)	(0.01)	(0.01)
N child	0.05^{***}	0.05^{***}	0.05^{**}	0.05^{***}
	(0.02)	(0.02)	(0.02)	(0.02)
BIC	2415.89	2417.67	2399.43	2405.16
Log Likelihood	-1144.56	-1145.45	-1136.37	-1139.23
Num. obs.	1731	1731	1724	1724

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

individuals who reported any type of damage in their neighbourhood are more likely to support such peace talks. More interestingly, individuals with any type of education are

Table 2.10: Impact of violence on the support of launching rockets against Israel

	Own Dwelling Damage		Neighbourhood Damage	
	(1)	(2)	(3)	(4)
Damage All	-0.17^{***}		-0.20^{***}	
	(0.06)		(0.06)	
Damage Major		-0.39^{*}		-0.00
		(0.20)		(0.06)
Age	-0.00	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)
Gender (Female)	-0.01	-0.00	-0.01	-0.01
	(0.06)	(0.06)	(0.06)	(0.06)
Education (Elementary)	0.27^{**}	0.28^{**}	0.27^{**}	0.28^{**}
	(0.12)	(0.12)	(0.12)	(0.12)
Education (Basic)	0.38^{***}	0.38^{***}	0.37^{***}	0.38^{***}
	(0.11)	(0.11)	(0.11)	(0.11)
Education (Secondary)	0.30^{***}	0.31^{***}	0.30^{***}	0.30^{***}
	(0.11)	(0.11)	(0.11)	(0.11)
Education (Tertiary)	0.26^{**}	0.27^{**}	0.25^{**}	0.26^{**}
	(0.12)	(0.12)	(0.12)	(0.12)
Income Hamas	0.56^{***}	0.58^{***}	0.59^{***}	0.59^{***}
	(0.10)	(0.10)	(0.10)	(0.10)
Emotional Instability	-0.01^{**}	-0.01^{***}	-0.01^{**}	-0.01^{***}
	(0.00)	(0.00)	(0.00)	(0.00)
HH Size	0.04^{***}	0.04^{***}	0.04^{***}	0.04^{***}
	(0.01)	(0.01)	(0.01)	(0.01)
N child	-0.06^{***}	-0.06^{***}	-0.06^{***}	-0.06^{***}
	(0.02)	(0.02)	(0.02)	(0.02)
R^2	0.05	0.05	0.05	0.05
Adj. R^2	0.04	0.04	0.04	0.04
Num. obs.	1740	1740	1735	1735

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

likely to have negative views on the peace talks in comparison to individuals with no education. As expected, households who received income from Hamas after the war are less likely to support the peace talks. Given the clear demarcation between supporters of Fatah and Hamas and their opinion regarding the peace talks as shown descriptively in table 2.2, the results do not provide novel insights.

The robustness of these results are further reconfirmed when using the third question regarding the support of halting rocket launching into Israel. Given that the variable has a clear order of preferences, it is treated as a continuous variable and OLS estimations are performed. The estimation uses the same specification and controls as before. The results of table 2.10 show that individuals who where exposed directly to damage from the war are more likely to oppose rocket launching. The coefficients for the major and any damage are both significant at the 10% and 5% levels, respectively. These results reiterate the general repulsion of supporting of extremist parties and views for the affected individuals.

2.8 Conclusion

This chapter analysed the impact of violence exposure on the electoral preferences of individuals in the Gaza Strip. Literature stemming out from the region and elsewhere have consistently found that individuals exposed to violence are more likely to harbour radical views, either through the support of extreme parties (such as Hamas) or of extremist actions. The chapter uses various indicators of violence exposure, measured directly through the destruction of individuals owns homes, and indirectly through destruction of other dwellings in the their neighbourhoods. Violent exposure in surveys has a number of advantages over other proxies, as it allows a clear attribution of violence exposure to individuals (Brück et al., 2016).

Violent exposure is shown to have an immediate de-radicalisation effect under various specifications and robustness checks. However, the impact disappears when prior electoral preferences are interacted with the violence exposure indicator. In other words, individuals who already support extremist parties will not switch their preferences to support less extreme ones, while moderates continue supporting moderate parties. Therefore, exposure of violence has a polarisation effect in the Gaza Strip after war. To that end, further research should focus on the long-term impacts of violence exposure to understand if such de-radicalisation and polarisation shock is only short-lived or is sustained in the long-run.

Chapter 3

Media Bias in Measuring Conflict: Evidence from Syria

3.1 Introduction

The use of sub-national geo-coded conflict event data in studying violence has surged in the past two decades (Buhaug and Gates, 2002; Kalyvas, 2005; Gleditsch et al., 2014, among others). Despite the growing availability and use of conflict event datasets, and its benefit over macro level data, a number of analytical and conceptual challenges remain unresolved (Eck, 2012). The most pressing issue is related to the reliability of the sources used in the selection process of events. Existing datasets rely primarily on secondary sources (e.g., media outlets and press agencies) for extracting information on conflict incidence (Sundberg and Melander, 2013; Raleigh et al., 2010). In order for an event to be included in a given database, it needs to pass through a long chain of information process. First of all, an event has to be observed and reported. Second, media sources then have to publish the incident based on various editorial decisions. Third, the database managers then have to include this particular media source in their pool of resources. Last, coders then transform the news into data points and classify them based on the relevant information available in the media reports (e.g., location, time, actors), before entering them into the final database. Therefore, selection or inclusion problems can arise at various stages of this data generation pipeline. For example, in the initial stages, biases in news reporting, either intentionally or inadvertently (due to lack of sufficient information), can have a myriad effect on the quality and accuracy of event data. This is true in the coding stage as well. Insufficient training of the coders can lead to low inter-coder reliability; and the lack of a unified definition of a violent event among scientists affects the type and number of events that are included in the final database (Eck, 2012). These challenges subsequently hinder the development of meaningful measures of violence and plague the analysis.

Media biases in news reporting has been explored comprehensively in the political science literature (Groseclose and Milyo, 2005, among others). This strand of literature has focused mainly on the causes of media bias in reporting. For example, reporting biases can be driven from the supply-side, associated with the editorial decisions made by a news agency relative to its political ideology (Iyengar and Hahn, 2009); or the demand-side by maximizing profit through slanting news content to cater to the beliefs of their viewership (Mullainathan and Shleifer, 2005). However, fewer studies have examined the consequences of such reporting biases, particularly in regards to the generation of conflict event data. Despite the widespread use of media reports in the collection of conflict event data, only just recently researchers started using rigorous quantitative analysis to determine the implications of such biases (Baliki, 2014; Zhukov and Baum, 2016; Weidmann, 2016). For example, Weidmann (2016) compares events from the conflict event dataset ACLED with a geo-referenced military dataset leaked by WikiLeaks in Afghanistan, while Zhukov and Baum (2016) examine an actor-based coverage of multiple news sources in the Ukraine. Both of these two studies find an undeniable evidence of the impact of the selected media sources on the results. Nevertheless, the literature is still in its infancy, and there are no studies that comprehensively match events across media sources at high spatial and temporal precision levels to determine the prevalence and magnitude of media bias, or offer viable remedies to it. Based on these knowledge gaps in the literature, two crucial questions arise:

How consistent and reliable is the use of media for the generation of conflict event data? And are there any methodological remedies available to ensure a more robust and consistent data generation processes?

The objective of this chapter is to answer these two questions. This is done systematically as follows: First, by quantitatively studying the reliability of media sources in providing unbiased and adequate event data. Second, by introducing an innovative methodology for the collection of conflict event data that does not rely on secondary media sources.

In the first part of this chapter, I analyse the extent and magnitude of existing biases in reports from media sources, both in regards to the selection and description of violent conflict events. The selection bias is determined by assessing the magnitude in under-reporting violent events by the media, as well as their geographic coverage of the conflict. While the description bias is measured via actor-specific propensities in reporting. The analysis uses a new micro-level event dataset on the Syrian war, which was collected as part of the pilot project DISC (The Documentation Initiative for the Syrian Conflict)[12]. DISC collected and coded conflict event data from five international and local secondary media sources (mainly via online published news reports) for a randomly sampled time period (60 days worth of news content) in order to extract and quantify violent and political information on the Syrian war. Moreover, the DISC codebook used in the coding process is built closely in line with existing event databases to ensure its comparability, yet also slightly augmented to capture context-specific information to the Syrian war.

The conflict in Syria acts as a viable example of the complex and severe nature of civil wars. After five years of an on-going multi-stage conflict, many remain sceptical that a viable way to bring peace will be found. The conflict has fractured the country into various mini-states that are controlled by multiple actors, and has attracted a number of international players into the conflict, both directly and indirectly (See section 3.3 for a detailed overview of the Syrian war).

The reliance on information gathered from multiple sources during the same period permits a fine-grained temporal and spatial matching of event data across these sources. Based on the identified matches, I then estimate the total number of observations, including unobserved (or unreported) events using capture-recapture models. This estimated abundance size is in turn used to quantify the magnitude of existing selection biases for each media source. The results reveal a wide gap of event selection. First, all of the five selected media sources report, at most, less than 10% of the actual number of violent events that took place in that pre-specified time period. The detection rates vary from one source to another, yet they are all significant at the 5% level. Second, media sources report violent events selectively in specific areas within Syria. By comparing the

---

[12] *DISC was a joint initiative between the Department of International Development at the London School of Economic and Political Sciences (LSE), and the Department of Development and Security at the German Institute for Economic Research (DIW - Berlin). The project was funded by the Security in Transition (SiT) and Training and Mobility Network for the Economic Analysis of Conflict (TAMNEAC) projects*

spatial spread of the reported violence for each source to a simulated spatial distribution of the estimated sample, a notable clustering effect in reporting is evident. Moreover, investigations on the reporting of actors between the matched events highlight further the depth of the biases. To that end, the results uncover significant shortcomings in the use of secondary media source for compiling conflict event data, as biases affect not just the number of events reported, but also their geographic allocation and description. This raises concerns regarding the reliability in the use of a media-based conflict event data for drawing sound analytical inferences, particularly in high intensity conflicts like in Syria.

The second part of this chapter proposes a remedy to mitigate the unreliability of secondary sources. This mainly includes the use of a novel methodology for generating conflict event data, namely crowdseeding. The on-going project "SCAPE" (Syria Conflict And Peace Events) collects data without the reliance on media reports, excluding the middleman completely. Instead, this data collection approach works with a number of selected reporters (*seeds*) living in various parts of Syria. The trained reporters enter data on violence and peace events happening in their vicinity on a daily basis into a pre-designed web-based platform. The core group of correspondents form the direct primary data source. This approach has not been implemented in conflict event data generation before at this scale (see Van der Windt and Humphreys (2016)), although other methods such as crowd-mapping and crowd-sourcing have been in use to monitor violence for a several years.[13]

The advantages of crowdseeding are outlined thoroughly in light of the consistent media biases found in conventional event data collection methods. To summarize, conflict event data based on secondary media sources retains an important place in the science. With the growth of online media sources, a vast number of media reports and resources are easily accessible to researchers. Nevertheless, the transformation of the reports into

13 *See Syria Tracker for an example, http://www.humanitariantracker.org/syria-tracker*

useful quantitative data is both time-consuming and inefficient. These shortcomings are driven by the continuing reliance on human-coding and the inefficiency of machine-coding. Moreover, and most importantly, secondary media resources are found to be plagued with biases both in terms of the selection of news and its conformity to facts, as will be seen in the first part of this chapter. In contrast, crowdseeding data reduces part of the biases inflicted by secondary reporting, and provides a good resource for modelling and shaping the type of data and the methodologies in use. Crowdseeding is still at its experimental stages. Its slow uptake and use in gathering conflict event data is plausibly due to the high risk surrounding the safety of the correspondents working on the ground, and its long-term sustainability in the provision of adequate information in protracted conflicts.

The rest of the chapter is structured as follows: Section 3.2 describes the concept of sub-national conflict event data, highlighting both its advantages over national level data, and its limitations, particularly in regards to their use of secondary media sources. Section 3.3 summarises the main events of the Syrian war to date. Section 3.4 introduces the dataset DISC in a detailed manner. Sections 3.5 and 3.6 describe the methodology used and the results on media biases in news reporting on conflict, respectively. The introduction of crowdseeding and the discussion around its merits are reserved for section 3.7. The last section concludes this chapter.

3.2 Literature Review

The number of studies that use disaggregated geographical and spatial datasets of conflict have increased remarkably in the past decade, which has stemmed from the significant work of, e.g., Stoll (1993), that looks at the causes and dynamics of conflict at disaggregated levels. In this research (e.g., Braithwaite and Johnson, 2012; Condra and Shapiro, 2012; Kalyvas, 2008; Linke et al., 2012; Lyall, 2009; Wood, 2010, among others), the causes of conflict can be traced to nuanced and multifaceted interactions between conflict players, as well as socio-economic, political, grievance and other intervening factors.

The surge in this research is indicative of three trends in the study of violent conflict and civil wars. First, social science researchers, and particularly economists, are increasingly using large datasets and econometric methods to investigate the subject (Collier and Hoeffler, 2004; Besley and Persson, 2009; Blattman and Miguel, 2010; Brück et al., 2016; Justino et al., 2013, 2016; Miguel et al., 2004). Second, there is a move to more micro-level and within-country analysis, as comparisons between countries have routinely failed to come up with robust explanations for violence. Third, the increasing availability of geo-coded event-based datasets of conflict and violence which permits such analysis at the micro-level (e.g. the Armed Conflict Location and Event Data Project - ACLED (Raleigh et al., 2010) and the Uppsala Conflict Data Program (UCDP) Geo-referenced Event Data - UCDP-GED (Sundberg and Melander, 2013)).

Using event-based data, researchers have been better able to exploit natural experiments and regional variations in violence to better understand the causes and consequences of conflict (see in regards to: economic behaviour (Bozzoli et al., 2013; Voors et al., 2012), health and nutrition (Minoiu and Shemyakina, 2014; Tranchant et al., 2014), political participation and voting (Blattman, 2009; Trelles and Carreras, 2012)), as well as allowing more accurate modelling of conflicts themselves (Fjelde and Nilsson, 2012; Metternich et al., 2013; Ferguson, 2015).

The following section highlights two important and relevant conceptual and empirical aspects in that regard: First, on the advantages in the use of geographical and temporal event data *vis-á-vis* macro-level indicators in the analysis of violent conflict. Second, on the current state of micro-level conflict event data and their limitations, particularly highlighting issues around biases in their use of secondary media sources.

3.2.1 From Macro to Micro

The shift from national to sub-national levels of analysis of violent conflict is owed to the conceptual and methodological limitations of country-level data. National-level analysis of civil conflict relies on two strong assumptions. First, conflict actors (the government in most cases) are equally present across all territories of a country. Second, aggregate country conflict indicators are capable of capturing the full spectrum of the internal variations of conflicts. These two assumptions are by design flawed. Most of the internal conflicts that broke out in the last five decades were concentrated in small geographic pockets, and did not spread across all territories of a country simultaneously. Hence the representation of a whole country as "in conflict" condones the fact that most of the territories are actually "in peace". Even attempts to disaggregate the conflict into administrative or political zones within a country does not consistently and rightfully account for the actual diffusion of the violence (Buhaug and Rød, 2006).

Consider the different cases of Syria and Colombia as an example. In Syria, there are currently more than a dozen of state and non-state actors involved in the conflict, who control various mini-state territories inside of the country. The intensity of the violence in Syria varies considerably both temporally and geographically in a given year. On the other hand, the Colombian conflict has been contained in specific geographic areas and involved mainly two distinct warring actors. Therefore, undertaking a cross-country analysis where Syria and Colombia carry the same weight in the data for a given year is not just falsely representative of the actual events, but also generates consistently biased and misleading outcomes.

In a nutshell, the use of national conflict indicators is inconsistent for the following conceptual and methodological reasons. First, fighting is likely to take place in small geographic areas and is unlikely to spread across the whole country. Second, internal fighting does not comply with administrative or country boarders, hence any spatially political and administrative representation of the violence is misleading. Third, internal conflicts usually involve more than one or two actors. Fourth, a yearly measure of onset of violence is not sufficient to capture the actual intensity and variability of the fighting. Fifth, temporal differences within a country are endogenously determined by the spread and intensity of violence. Hence cross-country fixed effect models are insufficient to account for such a variation. Sixth, the correlates and patterns of internal conflicts are heterogeneous across locations, and any attempt of aggregation could lead to inconsistent and biased outcomes. Seventh, aggregated national indicators typically ignore any contextual and attributable information of internal conflicts, possibly biasing causal inferences.

The shift towards the micro-level analysis of violent conflict and the increased use of disaggregated data provides fine-grained insights on the location, time periods, and actors, which was deemed impossible with the use of national level data. Despite the limitations arising in externally validating the results coming out of micro-level studies (Justino, 2009), there is a clear advantage in using sub-national event data in pushing the frontiers of conflict and peace research, as well as providing better answers on the causes and impacts of violence.

3.2.2 Current State of Micro-level Data

An event-based database is a way to operationalise existing media for new types of disaggregated spatial and temporal analysis. Two databases are widely used in the conflict literature, The Uppsala Conflict Data Program (UCDP-GED) and the Armed Conflict Location and Event Data (ACLED). UCDP was established in the 1980s, and since has refined its design and definition of armed conflict. The dataset covers most armed conflicts across the world. ACLED, on the other hand, mainly focuses on African (and recently South Asian) armed conflicts. Furthermore, both datasets are limited to violent events, and exclude local political and peace events such as agreements and peace negotiations. Given the importance of peace events in understanding the spread and end of violence, it is vital to include such local non-violent conflict events in the datasets.

These conflict event datasets are based on ex-ante conceptually developed “codebooks”. Such codebooks aim to direct coders on what violent events to include or exclude in the final database. UCDP-GED only defines, and consequently codes, an event as violent if it resulted in at least one death (previously 25 battle-deaths). Naturally, the larger geographic coverage of the UCDP-GED dataset comes at the expense of lenient coding rules. The imposition of strict inclusion restrictions of an events ensures comparability across conflicts from various regions. ACLED, in contrast, uses a less stringent inclusion restriction, and codes any type of event related to violence independent of its outcomes. This leads to a higher number of observations per conflict. Therefore, the disagreement on the use of a unified definition of an "event" can already have a notable impact on the significance of the results (Eck, 2012).

There are both conceptual and analytical disadvantages in the definitions applied in both datasets. ACLED does indeed provide a wider range of non-lethal conflict events. The large pool of events enables end-users to flexibly and selectively include events to fit their analytical needs. Yet, the lack of definitive inclusion restrictions on event types presumes that all events carry the same weight in the analysis - which can be problematic if event classifications by outcome are also disregarded. In a similar fashion, neglecting the importance some non-lethal events might carry - both economically (destruction of infrastructure) and socio-politically (displacement of individuals), UCDP-GED is likely to be insufficient in providing insight of singular conflicts. Moreover, the disagreement extends beyond mere definitions and the inclusion criteria. This includes: defining the actors involved (armed actors, civilians, etc), the level of geographic precision (points, grids, etc), and the outcomes of the violence (deaths, injuries, destruction, displacement, etc).

The biggest challenge, however, lies in the use of secondary sources in the coding process. All conflict data initiatives use a similar data generation pipeline. Information is extracted from media reports, which have been and are largely still the main bank of information available to data coders. Then through the coding process, this information is transformed into quantitative units. Reports are seldom cross-checked by other sources, and few resources validate the data points through inter-coder reliability. Here lies a structural problem that requires more than a unified codebook to remedy. The use of secondary media sources is both systematically biased and insufficient in the provision of a comprehensive and reliable information. Yet alternative methods to date are scarce and still do not address the core issues related to the use of media sources. For exmaple, The Global Database of Events Language and Tone - GDELT (Leetaru and Schrodt, 2013) uses automated machine coding for content analysis of news articles. Despite its time efficiency and its advantages in reducing human coding errors, the dataset still relies on secondary media sources and does not provide adequate measures for the validity of the information it codes.

3.2.3 Bias in Media Reporting

Measurement errors occur based on two reporting problems: the selection bias and the description bias. This mainly stems from the fact that media coverage of conflicts are rarely objective in regards to both the selection of news and the description its content (Earl et al., 2004).

The Selection Problem: This type of bias occurs when media outlets undertake selected reporting. News agencies are prone to report on violent events that they deem important to their readers or viewers. That is, media serves primarily to satisfy the political beliefs of their viewers, and not as documentaries of war (Bocquier and Maupeu, 2005). Selective reporting can be either deliberate or inadvertent due to what is termed as reporting fatigue. Violent events that occur solely in a country or region where violence has for longer periods of time not been prevalent (e.g., the Paris terrorist attacks) tend to appear on the news more than incidents that occur in a long-lasting conflict (e.g., Syrian War).

On the other hand, the deliberate exclusion of certain news imposes a far more serious problem, as the non-random under-reporting of events is more likely to affect the significance of inferences (Wooldridge, 2010; Weidmann, 2016). This is especially evident in local news media. Therefore, despite the gain in additional information by including local news in the coding process (e.g., to reduce reporting fatigue), such sources generate systematic errors on the occurrence of certain violent events. For example, Drakos and Gofas (2006) show that under-reporting of terrorism incidents in media are correlated with regime types. Terrorism incidents in countries with government-controlled press and low levels of press freedom are systematically lower than that of democracies, which explains the high observation of terrorism rates in democratic countries. Such systematic exclusion of violent events is highly problematic for the analysis of violence at the micro-level. Moreover, the frequent reliance on domestic media as sources of topical information is likely to pass the contagion to the international media.

Yet both deliberate and random exclusion of certain violent events are highly interlinked. This can be closely related to the "Rashomon Effect" of viewer demand - viewers are interested in knowing about some events, but not others. If one media outlet reports continuously on certain type of events to influence the viewer or reader rates, then given budget, time, and space constraints, they are prone to ignore other type of events. Moreover, in order to increase viewing or reading rates, the description used in the reporting process is shaped to quell the interests of their consumers. This evidently affects reporting bias through the issue of veracity.

The Veracity Problem: this type of bias occurs due to the political and economic interests of certain media outlets. Media and news agencies describe violent events as they see fit for their end users. This phenomena strikes at the heart of press objectivity, independence, and trustworthiness. When reporting a certain event, each news outlet will have a different story to tell, deliberately infusing the bias in the classification of an event. This can be as minor as not providing actor names, or as major as exaggerating civilian casualties.

Significance of the Bias: From a statistical point of view, reporting biases in secondary media source would not pose an issue if they are non-systematic (random). In other words, if the bias is not correlated with other independent variables in the regression, then its safe to conclude that it should not impact the results significantly (Wooldridge, 2010). Yet, this proposed econometric solution only applies to measurement errors resulting from non-deliberate selection biases, but becomes ineffective with errors resulting from the systematic selection or veracity biases.

The study of the phenomena of reporting bias in news is not novel (Earl et al., 2004; Groseclose and Milyo, 2005, among others). Yet, most studies do not dwell on how reporting bias affects the development and accuracy of conflict event data. To date, only few studies have addressed the issue of reporting biases in the secondary media sources (Baliki, 2014), and systematically or empirically studied its plausible impact on the results (Weidmann, 2016; Zhukov and Baum, 2016). Weidmann (2016) proposes four possible methodological solutions to account for reporting bias. First, is to do nothing if the researcher can safely conclude that the bias in the conflict data is random. Second, to apply the capture-recapture approach for estimating the number of unreported events. This approach has been widely applied in the field of natural sciences to estimate unobservable animal populations, and more recently in estimating casualty counts (Ball et al., 2003; Price et al., 2015). To my knowledge, there are no studies that use this approach in estimating the unobserved events for conflict data. Hendrix and Salehyan (2015) apply this technique to estimate the number of protests from two news sources, and found a large variation in under-reporting of events. Third, to use Monte Carlo experiments for sensitivity analysis on possible biases. The simplest form of this technique uses large simulations, mostly at boundaries values of the expected bias, and check if it affects the significance of the result. Fourth, to correct statistically for potential errors dependent on the type of regression analysis. These propositions, depending on the task at hand, could be useful if the researcher has sufficient information on the nature and magnitude of the bias in the data. However, in the absence of such information, none of these propositions is a panacea to the fundamental issues in reporting bias, and in turn is unlikely to yield consistent outcomes.

Using DISC, this chapter aims to unveil the uncertainty on the extent of reporting biases in warfare. The dataset was collected from five international and local news sources on Syria for a period of 60 days. The study period is divided into 15 segments, each constituting 4 consecutive days of news reports. The results from the spatial and temporal analysis show that there are vast and substantial differences between the news sources, underlining both selectivity and veracity problems. I find that there exists a large number of under-reported events with only 10% detection rate at best. Moreover, I show that there is a non-random spatial heterogeneity in the reported events, as well as source interdependence between the media outlets. In other words, the exclusion of violent events from the media sources are systematic. Hence, simply ignoring the issues of bias in any regressions analysis using media-based conflict event data is likely produce inconsistent estimates and false inferences.

3.3 Context

The 6-year ongoing Syrian multi-actor war is one of the most intertwined and dynamic humanitarian crisis of the 21st century. More than 250,000 Syrians are estimated to have lost their lives since March 2011, and over a million have been injured. About 5 million Syrians have been displaced externally, and another 6.5 million internally. According to the latest estimates of UNOCHA, 13.5 million Syrians, including 6 million children, require immediate assistance.[14] Several cities and towns have been completely destroyed to rubble, and the Syrian economy has contracted by an estimated 40% in the last 5 years. There is no end sight for the cessation of the fighting in Syria, and the use of violence is still on the rise. These staggering figures are a result of a dynamic complex warfare which can be summarised into five major periods:

Period 1: Protests and the beginning of the armed insurgency (March 2011 - September 2011)

What started as peaceful demonstrations demanding for political and economic reform in Syria, in the wake of the Arab Spring of 2011, turned bloody quickly. Unorganized protests and demonstrations started gaining momentum and attracting civil participants, significantly after the the Friday prayers. The "iron fist" policy of the regime and the use of indiscriminate force against unarmed civilians resonated to the outside world. The opposition started shortly after organising its grassroot movement. Incidents of kidnapping and detention witnessed a sharp rise in the early days of the conflict, yet all

[14] *http://www.unocha.org/syrian-arab-republic/syria-country-profile/about-crisis*

the violent and non-violent oppressive measures implemented by the regime failed. The transformation of the political opposition into an armed forced began with the defections of army officers and soldiers into the opposition and the formation of the Free Syrian Army (FSA).

Period 2: Escalation and sectarian violence (October 2011 - July 2013)

The establishment of the opposition armed forces under the umbrella of the FSA changed the course of the conflict considerably. Skirmishes and clashes were reported on a regular basis in various areas around Syria, especially in the governorates of Idlib, Homs, and Daraa. The regime started using indiscriminate heavy artillery and shelling against rebel-held areas, which resulted in the large destruction of dwellings and infrastructure and death of civilians. The first attempt to halt the violence (Geneva I) on April 2012 was a major failure. There has been a consent that a transitional government body with full executive powers is needed. However, the future role of the Syrian president in this transitional government was not agreed on. The collapse of the peace plan resulted in the withdrawal of the UN mission from Syria.

The intensity of the violence in the second half of 2012 increased substantially amid the diplomatic stalemate. Clashes spread to the largest two cities in Syria - Damascus and Aleppo. Islamist opposition groups such as the Al-Nusra Front started to play an important role in the war, attracting high number of fighters to their ranks, and consequently began controlling territories. The number of actors fighting on both sides of the war increased considerably during this period, and outside armed actors such as Iran-backed Hezbollah (Lebanese) became also militarily active in Syria. At the end of this period, there was a clear confessional and religious demarcation between the various factions in Syria.

Period 3: A growing international concern (August 2013 - May 2014)

The chemical attacks in the Ghouta region in the Damascus suburbs that killed hundreds of civilians was the trigger that awakened the international community on the extent of the deterioration of the Syrian Crisis. Moreover, a new armed group, the Islamic State of Iraq and the Levant (ISIS) started being active in the northern part of Syria. In early 2014, the Geneva II peace conference organized by the UN and backed by major international countries aimed to bring the major warring factions to the table. No agreement was reached, however, on the nature of the political process and the transitional government. The use of violence continued, and rebels groups started fighting among themselves for political hegemony and control of territories.

Period 4: After the establishment of the Islamic State (IS) (June 2014 - August 2015)

ISIS seized the vacuum in power and the divisions within the opposition groups and established itself as a caliphate. Its power grew through seizing Mosul - the second most populous city in Iraq, as well as the eastern city of Raqqa in Syria. The establishment of the Islamic State shifted the dynamics of Syrian conflict to another level, through dissolving country borders and using unprecedented intense forms of violence. Meanwhile, Kurdish-controlled areas in the North and Eastern part of Syria also established themselves as self-governing, a stepping stone for the establishment of a Kurdish state.

Period 5: The Russian intervention (September 2015 - current)

After the failed talks during the UN assembly at the end of September 2015, Russia decided unilaterally to intervene in the Syrian conflict to eliminate extreme opposition forces. The Russian military air force became an actor in the conflict for the first time and the intervention assisted the regime in gaining control of areas previously lost to opposition forces. Moreover, Turkey also recently intervened directly in the war, seizing areas held by the Islamic state and the Kurdish forces. This period signifies the direct involvement of major countries in the war.

In summary, the Syrian war in the past six years has witnessed unprecedented use of violence across various territories in a relatively short period of time, the establishment and activity of a large number of armed actors, the dissolution of country borders, direct interference of regional and international countries in the war, and the continuous failure of international peace processes and agreements. Given all these elements, the Syrian case provides ample information on the complexity of modern day civil wars. Studying these dynamics and patterns can therefore generate useful knowledge and evidence on the nature of contemporary violent conflicts.

There are no conflict event datasets that capture the day to day violence in Syria. Outside academia, a number of organisations and initiatives do gather information on Syria. However, this information tends to take the form of casualty counts (e.g. Syrian Revolution Martyr Database and Violation Documentation Centre), daily reports (e.g. Syrian Human Rights Observatory) or archives of "case-files" of human rights violation (e.g. Centre for the Documentation of Violations in Syria, Syrian Human Rights Committee). These adequately convey the severity of the situation in Syria, but are unsuited for quantitative empirical analyses on the spatial aspect of violence. Given the lack of conflict event data on Syria, I collect data from various sources on the war. The data collection is introduced and described in detail in the next section.

3.4 Data - DISC

DISC's main objective is to provide a source of quantified information on political and violent actions in Syria at the micro-level. Event data was collected from various sources during the same time period to ensure event comparability across media outlets. The codebook was closely designed in similar fashion to that of ACLED and UCDP-GED datasets, especially in the use of violent event types and definitions. However, there have been slight variations in the structuring of the events in order to capture incidents and information unique to the Syrian conflict.

DISC is appropriate for both descriptive and analytical research, given the extensive information allotted for each event. Each observation is complemented with a number of temporal and spatial variables capturing information on the event date and location, as well as the time and location precision levels. Moreover, full details on the forms, means, and outcome of each violent event are recorded, as well as the actors involved. The next subsections describe the design of DISC in detail, the data collection process, and provide basic descriptives of violent and political events in Syria.

3.4.1 Design

The data collection process of DISC is based on a developed codebook for the data and the coding process (The full codebook is attached in appendix 4.7). The data coding guidelines provide definitions of the coded events - these include both violent and political events, actors involved, as well as the location and temporal precision. Data was collected originally from seven secondary sources during the same time frame. The secondary sources included both traditional and social media outlets. The main purpose of collecting data from various reports during the same period is to expand the source selection in order to gather as much information as feasible. Moreover, such data assists in the identification of unreported or misreported events between news sources - which will be at the core of the analysis of this chapter.

3.4.1.1 Media Sources

DISC is based on reports from two types of secondary sources: traditional and social media. 5 outlets were chosen from a large number of the traditional media outlets, two of which are in Arabic. The five source include: (1) Agence France Presse (AFP) - an English version of the International French based agency; (2) BBC Monitoring - UK; (3) Associated Press (AP) - USA; (4) The Syrian Arab News Agency (SANA) - the official Syrian governmental agency; and (5) Al-Arabiya - A UAE-based Saudi news network with detailed regional focus on the Middle East news. The choice of these five outlets was based on a number of factors. AFP and AP reflect different international coverage of the news from both European and American point of views. Although both agencies may rely on shared reporters and sources to get their news in Syria, mainly basing their information on wired news from Reuters, there remains a subtle difference on the events and type of news being reported. BBC Monitoring was chosen to reflect a novel reporting technique, which as its name suggests, monitors freely available and open media sources around the world. This large database of BBC Monitoring could potentially project higher news reports on violence *vis-á-vis* AP and AFP. SANA was chosen as the local news source from Syria. This especially assists in gathering more information on government-held areas and internal political actions. Al-Arabiya was chosen as the counterpart to SANA given that they cover the opposition-held areas. News agencies like Al-Jazeera (both Arabic and English) as well as Reuters were unfortunately not included in the study, mainly due to lack of financial resources.[15]

All traditional English and Arabic news sources were accessible through Nexis (previously LexisNexis) and the search for individual articles included the following keywords for the English sources: "Syria AND (violence OR conflict OR protest(s) OR kill(ing) OR rape OR terrorist OR arm(ed) OR rebel(s) OR refugee(s) OR displace(d))". The use of other keywords such as conspirators, saboteurs, battle, dead, fighting, clash, shelling, kidnapping, torture, destroyed, was also attempted. However, in- or excluding these keywords either adds very marginally to the selection or, in the case of fighting and dead, brings out substantial numbers of irrelevant reports. Using Nexis for the Arabic news sources proved to be more difficult. The use of keywords does not always retrieve all desired articles. After a failed number of search trials, the coders were asked to manually go through all the articles from the two Arabic outlets for the selected time frame of the study, and individually pick up the relevant articles.

Collecting data from social media was a daunting task. Data collections was attempted from two facebook pages - Monasiqoun (supports opposition groups) and SyriaTruth (supports the Government) - as well as a trial from an online news repository for the opposition networks called "Sham News". The information found on facebook pages was sporadic and unreliable, due to the change of style in reporting over time and the prevalence of lots of missing dates. Sham News started reporting daily round-ups for each governorate after 2013, which was easier to follow up. Yet the report dates did

[15] *The addition of both these news sources could provide larger number of events, which can increase the number of matches across sources, and hence reduces the confidence intervals in the estimation of the abundance size.*

not match the pre-specified time frame of the study, so they could not be included in the analysis. Even though we gathered data from social media, obtaining exact and precise information on the events for all dates was challenging. Hence, to avoid any misinterpretation of the results, I exclude data from the social media reports in the analysis of this chapter.

3.4.1.2 Type of Events

An event in DISC defined as follows:

> *"An occurrence that is related to violence and the political conflict in Syria; either an instance of the direct use of violence or other forcible actions, protests, or related internal and external political events."*

An instance is only coded as a single event if and only if: (i) it takes place at the same day; (ii) it takes place at the same location; (iii) it is of the same event type; and (iv) it involves at least two of the same actors throughout. Unless all these four conditions hold, an instance is coded as multiple events. The coded conflict events are divided into three main types: (A) violent events, (B) other forcible events, and (C) political and organisational events (Definitions and further information on events are described in detail in the appendix of Chapter 3).

Violent events are at the core of the dataset. Identifying an event as violent is a challenge given that all the aspects related to it need to be factored in. As described earlier, there has been no agreement on a universal definition of violence. Thus, taking into account both context and theory, a violent event is defined as follows:

> *"An event that is related to harmful force through the use of any manufactured arms or the use of unmanufactured arms in a manner that results in bodily harm to another person or destruction or damage to buildings or infrastructure."*

Moreover, each violent event is subdivided into three core components: (i) the form of violence, (ii) the tools used in executing the violence, and (iii) the outcomes resulting from the use of violence. The division of a violent event into these three components is unprecedented in conflict event datasets. Prior work attempted to differentiate between various forms of violence without success (Eck, 2012). The classification of a violent event in this manner eases the categorization of events without needing to impose a pre-defined inclusion restriction. This classification method does not deviate from the type of data available in other sources, but it takes into consideration the complexity in identifying various types of violence with their respective outcomes.

There are three forms of violent events in DISC: violence between armed actors, violence against civilians, and military offensives without a clear target. Tools of violence include a comprehensive list of both manufactured and non-manufactured weapons. Lastly, the outcomes of the violence include: death and injuries, differentiated by civilians and armed actors; destruction, differentiated by type of dwelling or facility; and territory gains (losses) by armed actors (More detailed definitions and descriptions of the types, tools, and outcomes of violent events are available in the appendix of Chapter 3).

Other forcible events include all other war-related events that are used in conflicts but do not fulfil the criterion to be classified as violent events. In other words, these events do not involve the use of violent action as defined above. We identify 6 types of forcible actions: (i) forced displacement; (ii) kidnapping and detainment; (iii) organized theft and looting; (iv) sexual violence; (v) road and border closures; and (vi) utility cuts.

Protests, Political and Organizational Events are all classified as non-violent conflict events. With events involving protests and demonstrations, coders have been instructed to divide the event into two in case violence was used against the protesters. Political events on the other hand, included all relevant internal and external political and diplomatic actions that took place in relation to the Syrian war. The types of political events included: agreements, meetings, and statements issued by relevant actors.

3.4.1.3 Time Frame

The time frame of DISC is from March 2011 (start of the uprising) until June 2013. Due to limited financial resources, 15 segments constituting 4 consecutive days were randomly selected for every second month (approximately) within this time frame. In total, DISC contains information from a sample of 60 days worth of news reports. The choice of the 4-day segment was primarily based on providing information from various periods of the conflict. This allows the examination of reporting variation between different time periods in the life-time of the conflict. One disadvantage of such a dichotomous sampling technique is that it does not account for time differences between the occurrence of the events and the day of the report. Report dates do not necessarily include day-to-day information of the events that took place on those exact days. In some cases, coders found that some reports included events that occurred weeks and months earlier, while others included news briefs of the major events that took place in the past months. All events that could be detected in the report were coded nevertheless, specifying both report and event date separately.

In order to ensure comparability between sources, the final conflict event database limits the days of events to that of the reports. This helps reduce noise and inconsistency generated from various data points that are coded from news reports which fall beyond the specified time frame. Moreover, given that news is usually reported one day ex-post, events that happened one day before each segment are included in the final dataset. For robustness purposes, I also include all events that took place between the last days of each segment and the first day of the consecutive one. Random sampling media-based conflict event data is a challenging task, as the sample generation uses dates of reports, which are not necessary in line with the dates of events from these reports.

Table 3.1 lists the time segments of the chosen report dates, and the percentage of the events reported on those exact days. Events are divided between political and violent events in the table. The last two columns include also events that took place one day

Table 3.1: DISC's time frame: variation between report and event dates

		(1)	(2)	(3)	(4)
Segment	Report Dates (inclusive)	% violent events		% political events	
1	April 7-10, 2011	51.61	74.19	43.04	56.96
2	June 9-12, 2011	38.00	62.00	86.25	97.50
3	August 3-6, 2011	48.84	88.37	65.82	98.73
4	October 7-10, 2011	75.51	97.96	76.25	93.75
5	December 5-8, 2011	65.71	91.43	79.71	95.65
6	February 2-5, 2012	67.44	100.00	81.44	98.97
7	March 27-30, 2012	32.88	64.38	64.08	91.26
8	May 10-13, 2012	80.70	100.00	88.60	98.25
9	June 27-30, 2012	49.00	61.00	56.98	76.74
10	August 2-5, 2012	60.19	76.70	66.00	96.00
11	September 15-18, 2012	79.70	85.71	61.68	89.72
12	November 21-24, 2012	39.00	39.00	89.47	94.74
13	January 6-9, 2013	85.71	85.71	86.67	100.00
14	March 15-18, 2013	72.22	93.06	69.23	94.23
15	May 7-10, 2013	41.67	70.83	65.38	96.15

Note: *columns (1) and (3) present the percentages of events, where the date of event is exactly equal to the date of the report ($date_{event} = date_{report}$). Columns (2) and (4) present the percentage of events, where the date of event is exactly equal or one day earlier than the date of the report ($date_{event} = date_{report}$ or $date_{event} = date_{report} - 1$).*

before each segment. As can been clearly seen from the results of the table, there are no obvious consistent temporal trends in the style of reporting for either violent or political events. The selected list of news agencies did not increase nor decrease their reporting precision with time. Including "the day before" as can be seen in column (2) and (4) for violent and political events, respectively, increases the percentage of events dates falling within the reporting period considerably. For example, in segment number six, about 33% of the events reported were from the 1st of February, one day earlier than the pre-specified period. Moreover, and as expected, political events were more likely to be reported immediately and in shorter time periods, as they are easier to capture than violent ones.

3.4.1.4 Geographic and Temporal Precision

Geographic and temporal precision levels are an indispensable aspects of micro-level conflict event data. Location data is preferably coded at the highest level of precision, which include the exact location coordinates. Time is another important aspect of conflict data given that it allows users to identify when the events took place. Given the difficulty for coding the exact time an event occurred, the lowest precision levels are days. In DISC the following geographic and temporal precision levels are used:

Geographic Precision: Syria is administratively divided into 16 governorates. Each governorate is then further divided into 3 lower administrative levels (districts, sub-districts, towns and cities). The capital cities for each governorate are further divided into municipalities and then into neighbourhoods. In order to better represent spatially the administrative divisions in Syria, five precision levels are chosen. Table 3.2 shows these precision levels, ordered from the least precise (governorate) to the most precise (neighbourhoods within a city). The number of units within each division is depicted by N. μ(Area) is the mean area calculated for each precision level. As seen from the values of levels μ(Area), levels 1 and 2 are imprecise, while level 3 are precise.

Table 3.2: Geographic precision levels in DISC

Precision Level		N	$\mu(Area)$	min	max
1	Governorates	16	13,433	116	43,204
2	Districts	61	3,083	116	30,801
3a	Cities, Towns and Villages	5251	0.97	0.01	114
3b	Municipalities (e.g., Damascus)	18	6.4	0.43	32.94
3c	Neighbourhoods (e.g., Damascus)	100	1.17	0.06	13.88

Note: *All Areas are in km^2, 3b and 3c only represent summaries of Damascus as an example.*

Depending on the geographic precision levels available for each event, the coordinates (longitude and latitude) are recorded in the final database. For events with precision levels 1 and 2, no coordinates were included, but all events reported at level 3 and below were geo-coded. If reports only included the name of the city (level 3a), then the centroid of the spatial polygon representing the city is chosen as the default location. The same holds true for the district and the neighbourhood within each city. This technique can be implemented for higher precision levels, but is likely to only generate noise data. The larger the spatial polygon, the more the error in geo-coding events, especially if data from all precision levels are included in the same analysis. The advantage of having higher precision levels at the district or governorate is highlighted when aggregating the data. In other words, all events with precision level 3 can be represented at level 2, and both levels can also be represented at level 1. Depending on the type of analysis, such aggregation may be useful as it offers more data points.

Temporal Precision: Violent events rarely take place within a one-day period. This leads to the phenomena of temporal interdependence (i.e., an event stretches over midnight). In this case, a new separate event is recorded for each day until the end duration of the period. If a source reports an instance with reference to a period longer than a day (e.g., "last week"), it is only coded as multiple events with daily temporal precision if the source makes clear that the instances occurred every day during that period. If not, then it is coded as one event with higher temporal precision. Moreover, in case that a temporal precision is higher than lowest level (day), events are coded on the lowest median day of the period indicated. For example, the second day is coded if the source mentions a period of three days; or the third day if source mentions a period of five days.

DISC specifies four temporal precision levels: (i) Daily precision of time where the exact date of the event is known. (ii) Imprecise time (2-6 days), where the exact date of the occurrence is unspecified. The period in which the event may have occurred spans more than 24 hours, though not longer than six days. (iii) Weekly precision of time, where the exact date of the occurrence is unspecified, and is indicated to a certain week. (iv)

Monthly precision of time, where the exact date is unspecified, and is indicated to a certain month.

3.4.1.5 Actors

Actors are divided into three groups: armed actors, political actors, and civilians. The Syrian war has witnessed an unprecedented surge in the number of actors that have organized themselves in various locations. Moreover, various regional and international countries have directly or indirectly played a role in the armed conflict. Given that in the first two years of the conflict (DISC's time frame) countries did not involve directly in the war, only active local armed actors are attributed to violent events, but countries can be attributed to external political events (A full list of all political and armed actors is presented in the appendix of Chapter 3).

3.4.2 The Coding Process

In DISC, we trained coders to enter the data in an online platform, which was built in line with DISC's codebook. Several training sessions were undertaken to check the skills of the coders and test for inter-coder reliability - this measures the extent to which independent coders evaluate the characteristics of an event and reach a similar conclusion (Tinsley and Weiss, 1975). First, similar reports were distributed to all coders in the English language, and they were asked to recode the events into the platform. The first test trial showed considerable differences between coders, which led us to undertake a thorough training. This mainly addressed the challenges in the identification and classification of violent events. Multiple examples of various violent event classification were distributed to coders for reference. A second trial was conducted, resulting in significantly higher inter-coder reliability. Different writing styles of reports were a major challenge to some coders. Yet the biggest obstacle was the unavailability of enough information in the reports to meet the criteria set in the codebook, reflecting the weakness in using media reports for coding conflict event data.

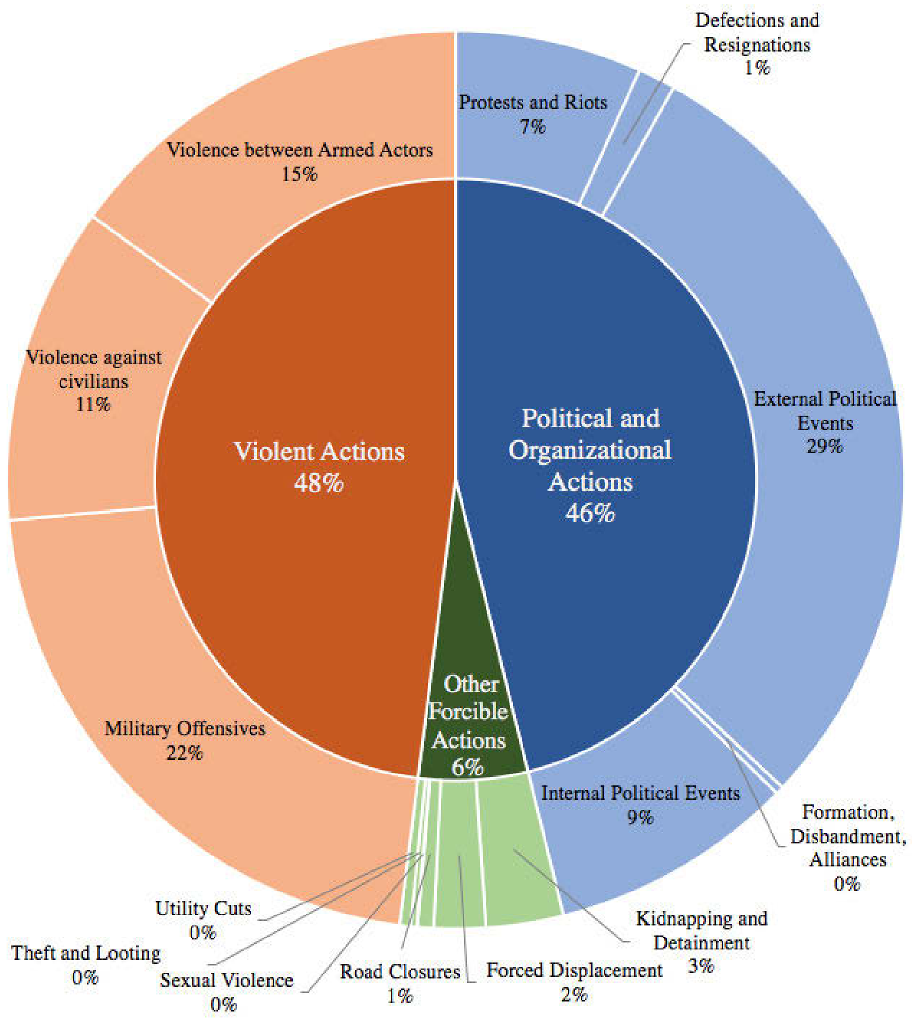
The form of the event is coded as it is described by the source media report. For each event only a singular form was selected. Coders read through the statements in the reports, assessed the content carefully, and then categorised the form of the violent event accordingly. Our definitions clearly state that a violent event should be coded as "violence between armed actors" only when skirmishes or combat actions are present, while coded as "violence against civilians" when the violence is intentionally targeted against civilians, which should be clearly stated in the report. The last category, "military offensive", was initially defined as violence targeted against a physical infrastructure, but also includes events that do not fall into the first two categories. That is, if the violence is one-sided and neither targeted against other organized armed actors or civilians. For example, the statement: "*The Syrian Regime fired heavy artillery on Aleppo*" cannot be truly categorised as violence against civilians or between armed actors since the targeted group is not explicitly mentioned. In this case we asked the coders to include the event under military offensives. Moreover, such statement does not clearly indicate if Aleppo is meant as the city or the governorate. For such cases, coders were asked to link the statement with the whole report context, and then apply their own judgement on what the reports actually is trying to convey. To ensure cross-validation and checking, the exact written statement from the report was included for each event. This allowed us to manually cross-check all data to screen for any errors arising from data entry.

3.4.3 Descriptive Results

Reporting Trends

DISC contains 2,798 events (observations) across all five sources. Most of these events are mainly violent (48%) or political (46%) actions. Figure 3.1 is a pie representation of all event forms grouped by the main three types: violent actions, political and organizational actions, and other forcible actions. 29% of all observations are categorized as "external political events", more than threefold of events categorized as "internal political events". Protests share 7% of all events, which have been given news attention particularly at the start of the war. The largest bulk of violence events, on the other hand, are categorized as “military offensives” with about 22%. This reflects the difficulty of categorizing violence reports from media sources as explicit violence between armed actors or targeted violence against civilians, which respectively represent 15% and 11% of the total number of events. Other forcible actions represent only about 6% of the total number of events. These are are mainly focused on incidents of displacement, kidnapping, and detainment. Other non-violent forms of conflict, such as utility cuts and theft for instance, represent less than 1% of total events found in the news reports.

Figure 3.1: Type of events in DISC - pie chart



Next a breakdown of the event types is undertaken by the 5 traditional media sources.

Figure 3.2: Share of reported events' type by source

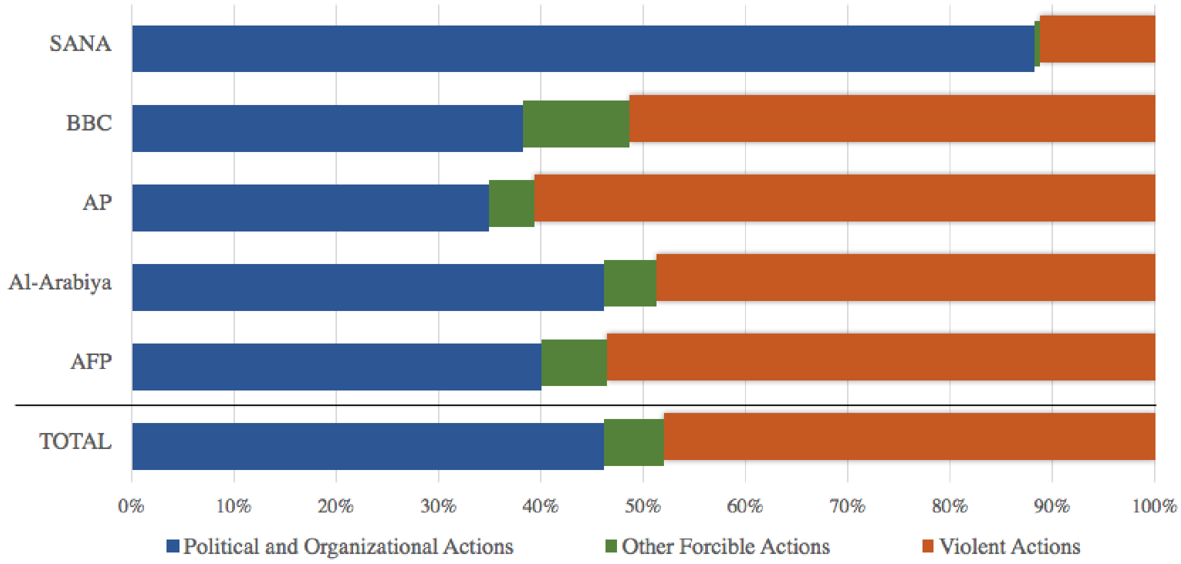
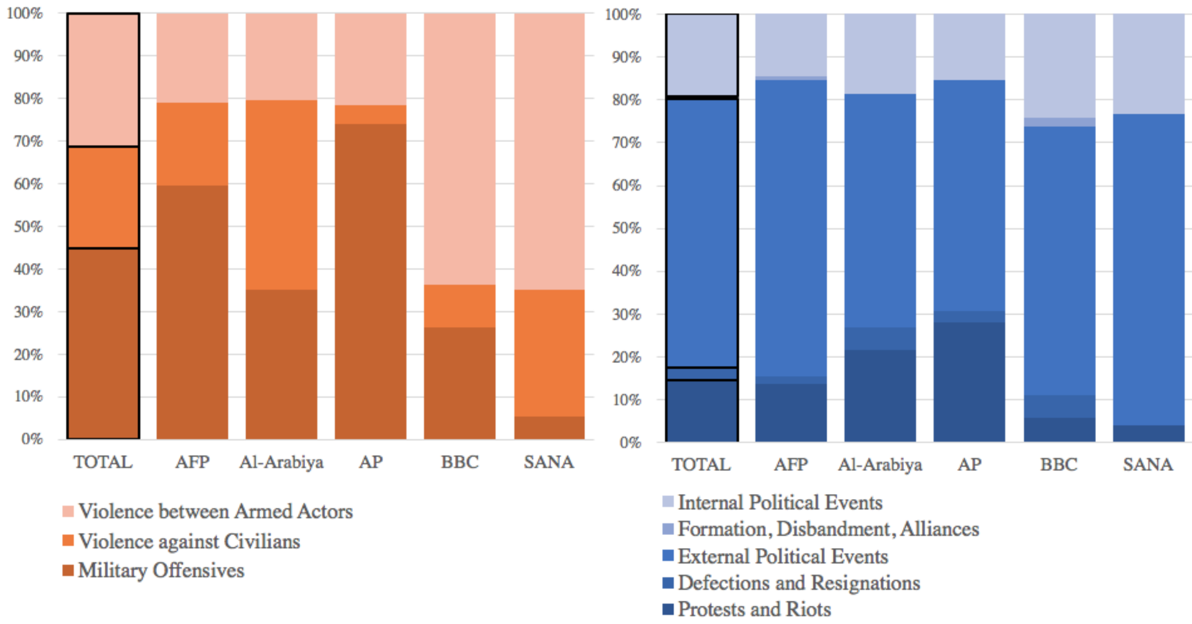


Figure 3.2 shows the percentage distribution of all three main events types by media source. The lowest bar of the figure depicts the total distribution of events as can be seen from inner section of the pie chart in figure 3.1. As expected, news reports from SANA have mainly focused on political events of the Syrian conflict in DISC's sample, and have devoted less attention to acts of violence. BBC Monitoring, Al-Arabiya, AFP, and AP on the other hand, retain a fairly uniform distribution of coverage between political and violent events. BBC Monitoring, however, has slightly devoted more news space to report on non-violent conflict incidents such as displacement and kidnapping.

These similarities in reporting trends extend to the coverage of political actions by source as can be seen from the rightward part of figure 3.3. Most the news sources focused on external political events related to Syria, and the distribution is uniform across all 5 media outlets. Al-Arabiya and the Associated Press reported more protests on average in comparison to the other sources about 20%-30% of the total number of political events. SANA on the other hand has barely devoted news space for protests (less than 5%). Moreover, fewer reporting space has been given to organizational actions such defections and formation of alliances, although their frequency was high during the first phases of the war.

Figure 3.3: Media coverage of violent and political events



Yet more importantly, how does this coverage trend vary dependent on the various forms of violence? The left part of figure 3.3 (in red) shows the share of violent events by dividing the violent actions further into its three forms: military offensives, violence targeted against civilians, and violence between actors. Here the variation between the sources is more evident, and several trends stick out. First, SANA having reported very few violent events in total, has focused mainly on violence between armed actors, and less so on military offensives (which are more likely to be perpetrated by the regime forces). Second, about 75% and 60% of reported violence by the Associated Press and Agence France Press, respectively, is classified as military offensives. International news agencies are less prone to specify the exact forms of violence without verification. Third, Al-Arabiya, which supports the opposition, classified the largest share of its reported violence events as ones targeted against civilians. All these differences are in line with the initial expectations on the political positions of the selected news agencies in regards to the civil war in Syria.

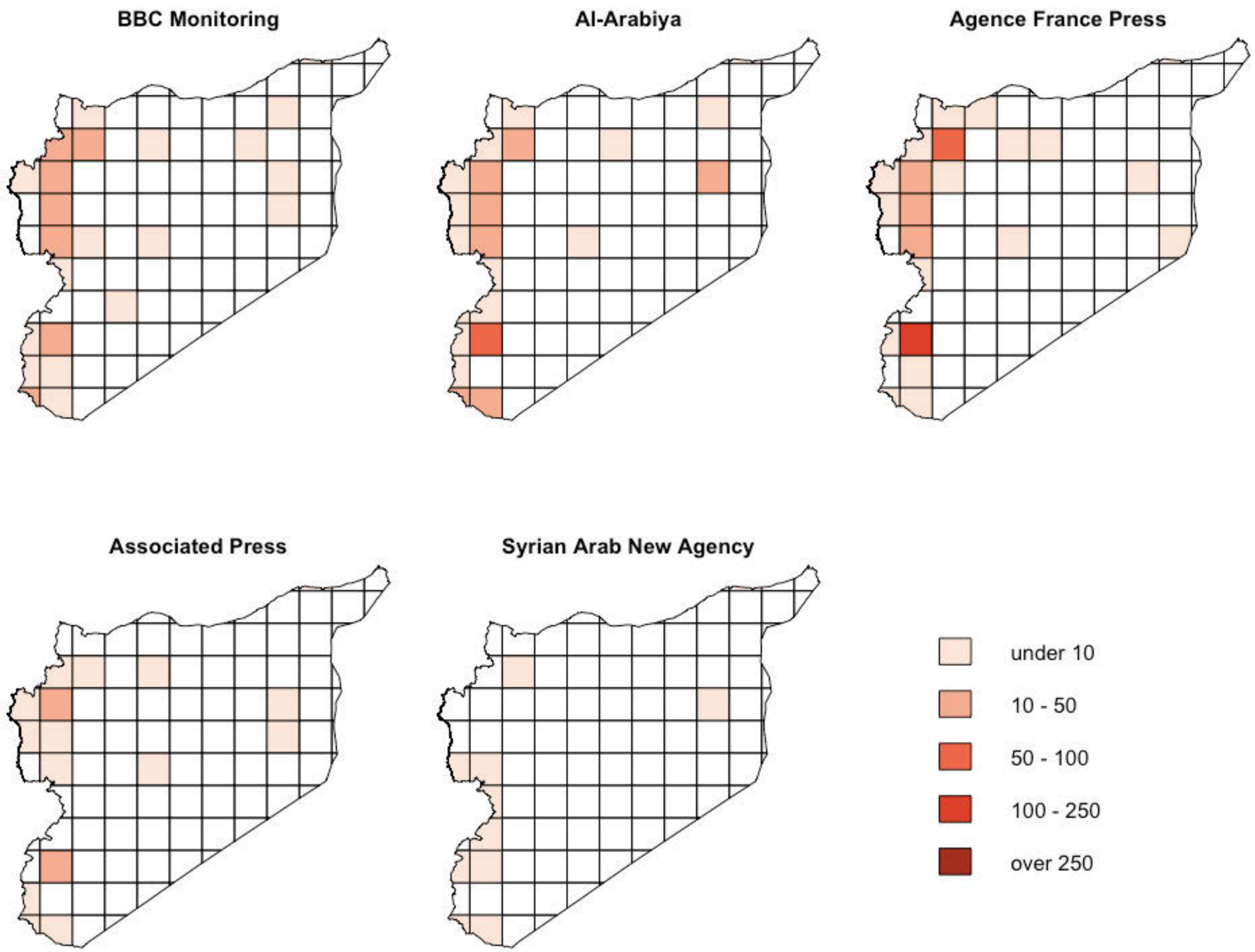
Location Accuracy

The total number of violence events in DISC is 1,298, varying considerably by media source: BBC Monitoring (300 events), AFP (397 events), AP(203 events), Al-Arabiya (361 events) and SANA (37 events). Moreover, not all violent observations specified high location precision which enables the geo-coding of these events. The total number of geo-coded violence incidents is 979 (about 75% of total number of violence events). The share of violent events that are reported at either the district or governorate level (i.e., without geographical coordinates) is uniform across most sources (between 16% and 24%), except for the Associated Press, where the share of non-geo-coded events is about 50%.

The difficulty in precisely narrowing down the location of an event is a major drawback of event data relying only on secondary media sources. Some of the events are excluded in the analysis given that media reports in certain times (here about 25%) only specify the location at the district or governorate level. Moreover, even if the town or city name is provided in the news report, the precise location of an event remains difficult to code accurately. Hence, the researcher has to make strong assumptions regarding the exact coordinates of such events. In order to reduce the likelihood of overestimating the bias given the variability of the spatial information, only the reduced geo-coded dataset is used in further analyses. The geo-coded dataset is constructed as follows: First, observations are dropped out if the spatial precision level is above the city or town. This is particularly important when using location as a constituent of the bias. Second, the coordinates of the polygon centroid of a city or town is chosen as the point of occurrence. This approach also applies for districts and neighbourhoods within a city.

Figure 3.4 uses the reduced geo-coded dataset to map the number of violence incidents by media source. The map representation uses a spatial grid system, which divides the geographic polygon of Syria into 50 km sided square grids. Border grid cells are

Figure 3.4: Gridded spatial distribution of violence incidence by media source



Note: *Each Grid cell has a side width of 50 km. Border grid cells are clipped to the border boundaries of Syria. Each grid cell contains the number of violence incidents reported at the lowest precision levels, and then aggregated to the grid level. All other event data at the geographical precision of the district or the governorate are not included in the analysis.*

truncated as appropriate to fit the projected polygon of Syria. The number of incidents for each grid cell is then calculated by simply counting violent events falling within the spatial boundaries of the grid cell. As can been seen from all five grid maps, there is an undeniable clustering of violence around major cities in Syria. This is particularly evident for the grid cells that contain Damascus, Aleppo, Homs and their immediate suburbs. The highest number of incidences reported by the AFP (more than 250 incidents) is mainly concentrated in the grid cell of Damascus and its suburbs. This number represents more than 64% of the violent events reported by this news agency. The spatial trend of reporting violence by Al-Arabiya is similar to that of AFP. BBC Monitoring, on the other hand, have covered larger areas in Syria, especially in the southern and eastern parts of the country. Despite these variation between the sources, violence events seem to cluster in specific areas of Syria.

Table 3.3 shows the Kendall correlation between the spatial grid units at both 25 and 50 km between all 5 media sources. Since there is a notable spatial variation between reported incidents at the grid level between sources and across Syria, a high number of grid cells are empty (with no violence incidents). This poses a challenge in providing a complete set of observations (at the grid level) to generate meaningful correlations. This problem can be remedied in two ways. First, by setting all grid units without any recoded incidents of to zero. The downside of such an approach, is that it increases the correlation values driven by all the zero units. Second, by alternatively setting all empty grids as missing observations, and running correlations within each grid. For example, if SANA reported incidences of violence in x grids, while Al-Arabiya reported incidences of violence in y grids, then the correlation only takes into consideration the intersection grid set $I : \{x \cap y\}$. Hence all non-zero observations of these two sources that do not fall in the intersection grid set are neglected.

As can be clearly seen from the results of table 3.3, the larger the grid dimensions are, the higher the correlation value of violence incidence is. Increasing the area of the grid cell from 25 squared km to 50 squared km steadily increases all pairwise correlations between sources. This trend is also consistent when setting the empty grids as zeros (lower part of the table) - which could have had an inverse effect, as the number of zero grid cells is reduced with larger grid sizes. This trend reiterates the problem of conflict research using units with high precision levels in their analysis, which is highly likely to lead to false positives.

Table 3.3: Spatial grid unit correlations between media sources

	50 km Grid Units					25 km Grid Units				
	Arab	BBCM	SANA	AFP	AP	Arab	BBCM	SANA	AFP	AP
Empty grids are recorded as missing										
Arab	1.00	0.69	0.18	0.78	0.64	1.00	0.49	0.05	0.49	0.39
BBCM		1.00	0.49	0.75	0.59		1.00	0.27	0.40	0.50
SANA			1.00	0.40	0.65			1.00	0.10	0.54
AFP				1.00	0.71				1.00	0.48
AP					1.00					1.00
Empty grids are recorded equal to Zero										
Arab	1.00	0.70	0.64	0.73	0.71	1.00	0.61	0.42	0.62	0.63
BBCM		1.00	0.51	0.72	0.73		1.00	0.39	0.63	0.58
SANA			1.00	0.50	0.45			1.00	0.34	0.39
AFP				1.00	0.71				1.00	0.63
AP					1.00					1.00

Note: *In the upper part of the table, grids that do not contain any violence incidence are recorded as missing, and only the intersection set between each sources is used to generate the correlation values. Hence, pairwise correlations with the Benferroni corrections are calculated between each two sources. The lower part of the table set all grids with no violence incidence to zero.*

Using only intersecting grids improves the correlation analysis, as it provides a better picture on the actual variations in reporting. Looking now at the results of the 25km grids: excluding the zero grids cells from the analysis can decrease the correlation (e.g., the correlation between SANA and Al-Arabiya drops from 0.42 to 0.05), or very seldom increase it, especially with low number of events (e.g., the correlation between SANA and AP increases from 0.39 to 0.54).

These correlation figures do not take into consideration time variability between incidents, and the arbitrary choice of the grid size substantially affects the results. Hence in order to provide a more rigorous assessment, precise geographical and temporal matching will be implemented. In summary, based on the initial descriptive analysis of the data, incidences of violence vary considerably across sources in terms of number of events reported, its classification, and its spatial distribution. These differences in reporting are likely to be driven from both selection and veracity biases, especially in regards to violent events. In the next sections of this chapter, I will undertake rigorous analysis to discern the extent of these biases by estimating their magnitude and spatial consistency.

3.5 Methodological Approach

As discussed earlier in the literature review section, media bias can be prevalent both in terms of inclusion of violence events in the reports (i.e., selection bias), or in terms of the description used in reporting the events (i.e., veracity bias). Hence in order to determine the prevalence and extent of both these biases, a set of different approaches is required. The following subsection will shed more light on the methodological design used to estimate each.

3.5.1 Measuring Selection Bias

In the absence of a reliable comprehensive "reference group", it is challenging to determine the actual number of violent events during the selected time frame. Thus, the measurement of the selection bias should be estimated from within the set of available sources. Simple count differences in the number of reported incidents between sources is an ineffective tool given that it neither accounts for the temporal and spatial attributes of these events, nor to the possibility of existing unobservables (events which are not reported in any of the selected sources). Hence, alternatively, the analysis uses a multi-stage rigorous approach to generate reliable estimates of the selection bias with the available tools and data. This approach entails the following sequential steps:

First of all, violence events will be matched on a one-to-one basis between all five sources. The criteria applied in the matching algorithm are primarily based on the location coordinates of events, their dates of occurrence, and their forms. This matching algorithm stems from the definition of an event as set out in the codebook of DISC. In addition, other matching techniques will be performed, which rely on more relaxed assumptions, and are used for robustness checks. Second, the capture-recapture approach is undertaken on the reduced matched matrix of events to uncover the best fitting model, which will be used to estimate the size of the abundance (i.e., the total number of observed and unobserved events). The selection bias can then be determined by calculating the detection rate for each news source. Third, the new estimated sample (which includes both observed and unobserved units) is used to generate Poisson Point Process (PPP) simulations of the spatial distribution of the events. The inhomogeneous random generation of the unobserved incidents are based on the geographic location of the observed data, as well as location drawn from population densities of Syria (i.e., covariates that influence the spread and intensity of violence). Finally, the observed spatial distributions of each source is then compared to the simulated total sample. Using Ripley L-functions with envelopes, the extent of the spatial intensity of the selection bias can be estimated. The methodological underpinnings for each of these four steps is described next in detail.

Event Matching

The one-to-one event matching is initially based on the specification of equation 3.1. The algorithm uses restriction criteria for event matching in line with the definition of a violent event as specified in the codebook. This approach guarantees the spatial, temporal, and categorical uniqueness of the matches. Thus, if i is an event from source I, and j is an event from source J then,

$$i == j \iff \begin{cases} coordinates_i = cooridnates_j \\ date_i = date_j \\ type_i = type_j \end{cases} \tag{3.1}$$

where *coordinates* depict the longitude and latitude of the events - i.e, the longitude and latitude of both observations i and j should be equal for a match to take place. *date* represents the specified date of when an event is reported to take pace; not the date of publishing the news report. *type* is the form violence as coded for i and j. These matching criteria are strictly restrictive. In order for a match to be recorded, all the three specifications should hold. The advantage of such a restrictive matching specification is that it does not allow for two occurrences from the same source to have similar criteria.

Remember that an event is coded as a singular event if it happens at the same exact place and time, is of the same type, and it involves at least one same actor throughout. Otherwise, the event was coded as two separate events within each source. For example, if a news source reported skirmishes between armed actors and violence targeted against civilians in the same location and at the same date, then two events were coded. Therefore, the specification of an event type in equation 3.1 is important to assure the uniqueness of the matched events. On the other hand, it is possible that source I reports an event as a military offensive, while source J reports the same event as violence against civilians (veracity bias). Hence, this matching specification can undercount the number of matches, which could in turn affect the estimates. The same principal holds for the location of an event. Similar events can, for example, be reported at the town level in source I, but at the governorate level in source J. In order to overcome and reduce any measurement errors resulting from the choice of the matching algorithm, robustness checks will be conducted using less restrictive criteria. First, by excluding the type of an event from the matching algorithm, as these can be driven by the description bias.

Second, by aggregating all locations to the governorate level.[16]

The matching algorithm generates a boolean matrix of 5 columns representing the sources, and n rows representing the observed (not captured/reported) number of events in the sample. 1's are replaced in each cell if the event i is reported by a source, and 0 otherwise.

Capture-Recapture

Capture-Recapture approaches have been originally used to count wildlife populations, and more recently their application has been expanded to estimate casualty counts in wars. The latter approach has been previously implemented in several contexts such as in Peru (Ball et al., 2003; Manrique-Vallier et al., 2013) and Syria (Price et al., 2015). The basic idea behind the capture recapture approach is to estimate the size of the population (or events in this case) in the absence of comprehensive observed data. A number of trials (or media sources in this case) are undertaken, where in the first trial, captured animals are marked and released back to the wild. In the second trial, the percentage of captured animals with marks are then used to estimate the population size of the area under study. The same principal applies with casualty counts, where a number of lists are used to match the names of the victims. A number of challenges arise from using such estimates, especially in trying to link variations in the count temporally and spatially. Such approach usually requires large amount of data to be properly disaggregated in both space and time in order to provide sound estimates (Landman and Gohdes, 2013).

Fortunately, this does not apply to conflict event data. Yet to my surprise such techniques have not been implemented before to estimate the number of unobserved violent events, apart from estimating protests (Hendrix and Salehyan, 2015). The large number of observations generated from the sample of conflict event data does not pose any notable

[16] *Matching events at the governorate levels allows the inclusion of the full sample of reported violent events, since the location does not depend on the geo-coded coordinates. Moreover, if we relax the matching criteria to lower precision levels, then the uniqueness of the one-to-one source matching does not hold, implying that more than one observation from a source I could match a singular observation from J.*

challenges in that regard, which allows to reliably use data both matched spatially and temporally.

The boolean matrix generated through the matching algorithm represents a closed capture recapture experiment, where n depicts the number of events, and S depicts the number of trials. Using Poisson log-linear models, I am able to estimate the total number of observed and unobserved events. However, there is a large number of models that can be applied in this context. Choosing the correct one can be a challenge, as one has to factor in a number of assumptions that describe the data best. These assumption involve the coverage and accuracy of the observed data, homogeneity, and source interdependence.

First, the "population" is assumed closed given that the trials are represented by sources on a closed time frame rather than temporal trials. Hence, the temporal invariability of the sample is sufficient to conclude that there are no major changes in event occurrence between one source or the other during the reported period. As a case of support, during the query, the selected news sources were found to update the same report throughout the day when more information was available. Coders have retrieved multiple updates of the same article, but only the latest version were used in the coding process. The assumption of a closed experiment simplifies the estimation of the models, as it is not vital to account for birth/death or migration corrections (as is the case with studying animal populations).

Second, "accuracy" involves mainly the use of a perfect match of the events, as well the existence of no false reports. The former condition is satisfied as explained in detail previous subsection. The latter condition, however, is harder to justify as it lies beyond the powers of the researcher in this case. Nevertheless, all five sources are renowned national and international outlets, that are less likely to fabricate news or make up events intentionally. The lower number of events in SANA also signifies that any changes to the reality might be driven by underreporting events. However, the assumption of no false events should be approached carefully despite the fact that there is no methodological way to correct for it.

Third, I assume that there is heterogeneity across sources. The homogeneity assumption requires that each event has an equal probability of being reported by each source. Given the nature of media reports on conflict and based on results from prior literature (see section 3.2), it is difficult to rely only on models built on the assumption of homogeneity. The good news is that analytical models accounting for unit-source heterogeneity are available.

Finally, interdependence of sources exists when media outlets get their news from each other, or from other joint banks of information. In other words, if source I reported an event, then it is more probable that source J will also report the same event. This is highly likely in our dataset, given that the news outlets are not mutually exclusive in obtaining their information on the Syrian war. Most international news agencies in the case of Syria either used similar secondary sources (e.g., Reuters) or have quoted each other. The difficulty in having independent reporters on the ground and the source competitiveness in providing on-going news information ("keeping up with Joneses"), is likely to increase inter-source citations. Hence, I hypothesize that the models using interactions between sources to account for interdependence are likely to perform better than ones that do not. The package *RCapture* in R allows to run models using GLM estimators that satisfy all possible combinations of these interactions, while controlling for the heterogeneity in the data. The best fitting model is then chosen from the large list of models based on the Bayesian Information Criterion (BIC).

Estimating the abundance size is straightforward in the case of only two independent sources as performed by Hendrix and Salehyan (2015), as the abundance parameter $\hat{N}$ can be easily approximated by using the century old Lincoln-Petersen estimator. However, when the number of sources increases (as is the case in the current dataset), it is required that generalized linear models using poisson distributions are used. The GLM analysis produces maximum likelihood estimates of the loglinear parameters, which is done stably through an iteratively least-squares algorithm (Baillargeon and Rivest, 2007). Then an estimate of the abundance $\hat{N}$ is derived from the loglinear parameters. In order to account for heterogeneity, three log-linear models are used: Chao, Poisson and Darroch (Chao, 1987; Darroch et al., 1993; Agresti, 1994; Rivest and Baillargeon, 2007). Chao applies a lower bound for the estimate, both with and without source effects which contains $S-2$ parameters. Alternatively, Poisson and Darroch models for heterogeneity are defined as $2^k - 1$ and $\frac{k^2}{2}$, respectively; where k is the number of captures.

Moreover, and in order to validate the abundance estimates of the log-linear models, especially under the conditions of heterogeneity and media source dependency, Markov Chain Monte Carlo (MCMC) simulations are conducted. The use of this method is recommended when the differences in log-linear models are marginal. Its main advantage is that it uses Bayesian non-parametric methods to estimate the size of a closed population based on the Dirichlet process mixtures (Dunson and Xing, 2009; Manrique-Vallier and Reiter, 2014). The MCMC algorithms are based on Gibbs sampling scheme to obtain samples from the posterior distribution of the model, including the estimates of the abundance size $\hat{N}$ (Manrique-Vallier, 2016).[17] Finally, using the estimate(s) of $\hat{N}$, the average probabilities of capturing an event for each source are calculated. These probabilities act as proxy for the detection rate, which in turn represents the desired magnitude of the selection bias for each source.

Poisson Point Processes

Using the abundance estimate $\hat{N}$ of the total number of observed and unobserved events, simulations are drawn from the intensity distribution of the data points to provide an accurate depiction of the spatial coverage of violence for each source. The number of events simulated are uniform across all sources (to facilitate comparison between sources) which are randomly distributed within the spatial polygon of Syria. Given the non-random nature of conflict event data, the assumption of Complete Spatial Randomness (CRS) is likely to be violated. Hence, the simulations generated from a homogeneous poisson dis-

[17] *The log-linear and the MCMC models are the two most widely used state-of-the-art estimates for abundance size in capture-recapture models for multiple sources ($S > 2$) with heterogeneity.*

tribution process become uninformative and will lead to biased estimates. Alternatively, I will use the inhomogeneous poisson process based on the spatial intensity distribution of the events for each source. This approach generates simulations taking into consideration the underlying distribution of observed events as well as other confounding factors that influence their distribution (in this case, the population density). In other words, the simulated data points depict the likelihood that a source S will report an unobserved event $\hat{E}$ in location L conditional on that an event E is already observed to be reported by S at L, and independently of S is likely to occur at L, i.e., $Pr(E|L) > 0$.

Choosing the exact underlying inhomogeneous spatial data generation mechanism is not a trivial task, particularly if the probabilities of capturing an event are low. Hence in order to reduce measurement errors, the simulated inhomogeneous random process must factor in the boundary distributions of the observed data and the covariates of violence. In other words, the distribution of the unobserved events is determined based on the locations in the observed data (observed distribution) and locations of population densities (fitted distribution of the covariates). Hence, the approximated spatial distribution of the unobserved events is generated by finding an optimal combination of these two distributions. Using the new point process with abundance $\hat{N}$, the Ripley K-function is estimated to measure the spatial bias for each source (Dixon, 2006), such that:

$$\hat{K}(d) = \frac{1}{\lambda^2 A} \sum_{i=1}^{n} \sum_{j \neq i}^{n} I_d(d_{ij})$$

where A is the polygon of Syria, λ is the intensity of events for each source, d is the distance from point i (j is all points other than i), and I_d is an indicator function that takes on a value of 0 if point j is not within distance d of i and a value of 1 otherwise.

Yet, the nature of the K-function makes it problematic to graphically discern differences between K_{theo} (which is based on the intensity of the second order point process) and $\hat{K}$ at lower values of d. In order to account for such limitation, it is advisable to transform the K- into an L-function as follows:

$$\hat{L}(d) = \sqrt{\frac{K(d)}{\pi}}$$

Envelopes of L_{theo} are generated based on Monte Carlo simulations using the second order inhomogeneous distribution of the unobserved events. This is necessary since simple comparisons of $\hat{L}(d)$ to L_{theo} are insufficient to determine the level of significance of dispersion or clustering. Thus, for each media source, if $\hat{L}(d)$ lies within the simulated envelope, then the media bias is significantly equal to zero at distance d . If $\hat{L}(d)$ lies above (below) the envelope, then there is a positive (negative) bias.

3.5.2 Measuring Veracity Bias

To determine the extent of veracity or description bias between sources, there are two ways forward. The first option is to use the subset of the matched data points and compare descriptively their event attributes (actors involved, outcomes, *etc*) one by one. This approach is descriptively very precise as it sheds light on how do various sources describe the same event differently. Its main drawback however, is that it ignores the probable selection bias between the sources, and it is only viable if all sources largely report similar events (as will be later shown that it is not necessary the case). Alternatively, one can examine the likelihood how a media source describes certain aspects of an event, conditional on that an event is reported in the first place. Although the precision of this approach is weaker than the first, its main advantage is that it uses all the data points available in the dataset to draw its predictive power. Moreover, this approach accounts for any plausible selection biases, as the propensities rely on simulations of the observed data. This facilitates predicting the extent of the veracity bias between sources without having to fall back on strict matching algorithms. Therefore, for measuring the description bias, I will use the latter approach. The average veracity bias is measured without explicitly accounting to either the temporal or spatial attributes of an event. This is done by using bivariate logit analysis on the pooled data, and then jointly predicting the expected conditional probability of reporting a certain description to an event (e.g., actors involved) for each selected media sources, while controlling and keeping other event attributes at their means. This approach relies on Monte Carlo simulations based on the observational data to jointly predict these propensities.

In this chapter, the focus will only be on the armed actors attributed by each source to an event. The armed actors are grouped into either government or opposition forces. Given that each event may have more than one actor attached to it, it is important to use bivariate logistic regression models to jointly predict the two binary dependent variables (Y_R, Y_O) where R and O denote Regime and Opposition forces, respectively. Hence for each event i there are four possible outcomes: $(Y_{iR} = 1, Y_{iO} = 1)$, $(Y_{iR} = 1, Y_{iO} = 0)$, $(Y_{iR} = 0, Y_{iO} = 1)$, $(Y_{iR} = 0, Y_{iO} = 0)$. The systematic components of the model are then equal to the predicted marginal probabilities $\pi_j = Pr(Y_j = 1)$ of each outcome and the odds ratio parameter ϕ (which captures the relationship between these outcomes and is defined as $\phi = \frac{\pi_{00}\pi_{01}}{\pi_{10}\pi_{11}}$). While, the stochastic component of the joint probability is $\pi_{RO} = Pr(Y_R = 1, Y_O = 1)$. Thus for $j = R, O$

$$\pi_j = \frac{1}{1 + exp(-x_j\beta_j)}$$

and

$$\phi = exp(x_\phi \beta_\phi)$$

Based on these parameters we can then determine the expected values of the model by drawing simulation for $\beta_R, \beta_O, and \beta_\phi$ from the sampling distribution and substituting them for the systematic components.

3.6 Main Results

3.6.1 Prevalence of Selection Bias

Table 3.4 shows the percentage of matched events across all five traditional media sources. The number of shared events between any two sources does not exceed 34%. The Associated Press (AP) has the highest number of events that are matched across other outlets. This could reflect the possibility that the AP relies on news sources from the AFP for example, due to the low number of events reported by the AP *vis-á-vis* the other three international news sources, the high percentage points do not seem likely to be outliers. On the contrary, only over 10% of the events reported in the AFP are covered in the AP.

Table 3.4: Event matching by source

		Comparative Source				
		BBCM	AFP	Al-Arabiya	AP	SANA
Base Source	BBCM	**230**	4.91%	6.60%	12.50%	0.10%
	AFP	6.96%	**326**	12.85%	32.69%	0.00%
	Al-Arabiya	8.26%	11.35%	**288**	22.12%	16.13%
	AP	5.65%	10.43%	7.99%	**104**	2.23%
	SANA	0.43%	0.00%	1.74%	0.96%	**31**

Note: *For example, 9.69% of events reported in BBCM are matched with ones reported in AFP, while 4.91% of events reported in AFP are matched with ones reported in BBCM.*

The most striking result is the low matching between BBC Monitoring and the rest of the news sources. Given that the BBC Monitoring relies mostly on other news sources in its reporting, a higher matching rate was anticipated. One should be careful in approaching the matching results at first sight. Given the short periods of reports dates chosen for the study, it is likely that some news agencies do no report the event until days they occur, and hence reducing the likelihood that the sources report the same events at the same dates. Taking a closer look at the data, the matched events that were mainly covered by most sources are the ones that have attracted immediate attention when they took place.

Table 3.5: Frequency statistics of the matching algorithm

	Unique	2 Sources	3 Sources	n_i	u_i
BBCM	193	25	12	230	230
AFP	260	43	23	326	269
Al-Arabiya	228	36	24	288	277
AP	54	27	23	104	54
SANA	26	3	2	31	26
Reported Events	761 (78%)	134 (14%)	84 (9%)	979	
Captured Events	761 (89%)	67 (8%)	28 (3%)		856

Note: *Unique is the number of units that is captured only once by one source. n_i is the number of events captured by source i, and u_i is the number of events captured for the first time by source i. The matching criteria includes location (latitude and longitude), date of event, and type of violent event. No events are captured by more than three sources, hence the columns 4 and 5 Sources are removed (all zeros).*

The total number of reported events in the subset dataset that includes only the geo-coded locations is 979. Based on the matching criteria of the specification 3.1, the number of captured events is 856. Table 3.5 shows the number of events captured by each source that is unique to the source or is shared by other news sources. First of all, the maximum number of sources reporting the same event is 3. Hence not even one event from the sample, based on the preferred matching criteria, was reported by four or all five sources.

n_i is the number of events captured by source i, and u_i is the number of events captured for the first time by source i. Apart from the Associated Press, most of the events reported by source are unique to this source. The captured events are equal to the reported events divided by the number of captures. 761 events were captured by one source only, 67 events by two sources, and only 28 events by three sources. In total, 89% of the captured events are then unique to one source. The low share of common captures and the prevalence in variations of the number of reported events across sources reflects the heterogeneous nature of the data. The prevalent heterogeneity in the data, and the large share of unique events, prompts the inclusion of interaction terms between media sources in estimating the abundance size from the capture recapture model (Baillargeon and Rivest, 2007). These interactions between the media sources account for the plausible interdependencies.

Next, using the matrix of captured events with $N = 856$, I then estimate the total number of observed and unobserved events (abundance). The estimations of $\hat{N}$ are initially drawn from various models which account for both heterogeneity and source variability. Table 3.6 shows the abundance estimates of these selected models with their standard errors, the goodness-of-fit criteria (deviance), as well as the AIC and BIC. $M0$ represents here the basic model which assumes that sources are static and homogeneous. Mt represents models that are only source variant, and Mh represents only heterogeneous models. Lastly, Mht represents models with both heterogeneity and source variation. In line with the descriptive statistics of the data, heterogeneous and source variant models perform much better than static and homogeneous models. Again, the homogeneity assumption requires that each event has an equal probability of being reported by a given source, without the need for the probabilities to be equal across all sources. Given that number of reported events by some sources (AFP and Al-Arabiya) is larger and more comprehensive than others (AP and SANA), than there is higher likelihood that the reported events are not driven by pure chance. Hence certain events which are reported by one source (e.g., AFP) are intrinsically and inherently different from events reported by another source (e.g. SANA).

From the upper part of the table *Mth Chao* retains the lowest BIC and has an abundance estimate of $\hat{N} = 3822$. Figure A.1 in the appendix shows the Pearson residuals of all

Table 3.6: **Abundance $\hat{N}$ estimates from various capture-recapture models**

	abundance	stderr	deviance	df	AIC	BIC
M0	2895.3	226.0	520.389	29	604.654	614.159
Mt	2682.8	205.1	114.978	25	207.243	235.756
Mh Chao	4087.1	446.7	480.247	28	566.511	580.768
Mh Poisson2	3983.4	450.6	505.736	28	592.000	606.257
Mh Darroch	8657.9	1974.6	494.531	28	580.796	595.053
Mth Chao	**3821.6**	412.4	68.010	24	162.275	**195.541**
Mth Poisson2	3859.0	432.2	95.767	24	190.032	223.298
Mth Darroch	9473.1	2224.9	82.659	24	176.924	210.190
Mth Chao[13,34,2,5]	**4085.1**	515.8	26.810	22	123.113	**165.884**
Mth Chao[34,1,2,5]	**4666.1**	582.8	35.408	23	129.712	**167.730**
Mth Chao[24,34,1,5]	**4880.4**	624.6	28.722	22	125.025	**167.796**

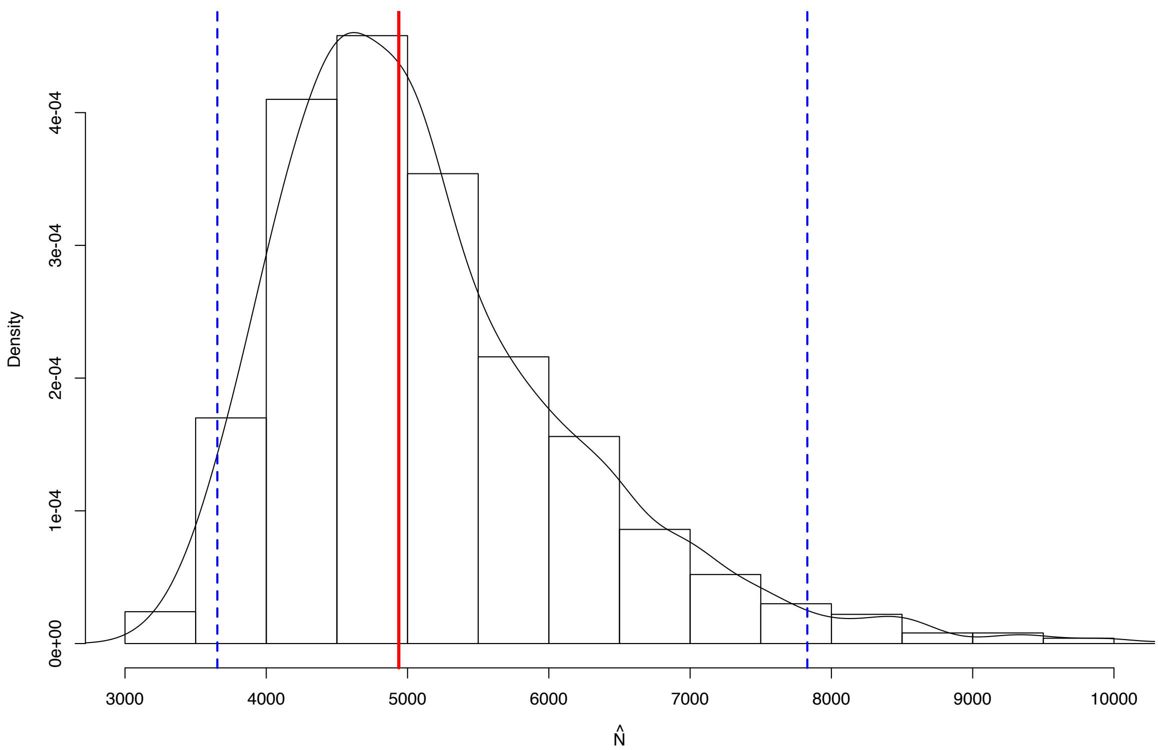
Note: *1=BBC; 2=Al-Arabiya; 3=AFP; 4=AP; 5=SANA. For example, [34,1,2,5] implies that there are only dependency between AP and AFP.*

the generalized linear model estimates. The Mth Chao estimate is also found to perform better than its counterparts Mth Poisson and Darroch, as its model residuals are the closest to zero.

Yet the initial estimate from the Chao model does not account for possible interdependencies between sources. Therefore, using only the Mth Chao estimates, I examine all possible dependencies between the media sources. The three models with the lowest BIC are listed in the lower part of table 3.6. The abundance estimates are higher in all three cases in comparison to the Chao model without controlling for these dependencies. Moreover, all three models perform better according to the AIC and BIC. The variation in the abundance estimates between the three models is quiet notable. Unsurprisingly, all three best performing models account for the source dependency between the AFP and AP. Yet they differ based on the following: The first accounts for the interaction between BBC Monitoring and AFP and yeilds an $\hat{N}_{Chao(1)} = 4085$ with a BIC of about 166. The second includes no other source interaction which yields an estimate $\hat{N}_{Chao(2)} = 4666$ with a BIC value of 168. Finally, the third best fitting model includes interactions between Al-Arabiya and the AP, which has an $\hat{N}_{Chao(3)} = 4880$ and a BIC of 168.

All three model have very similar BIC and AIC values, which makes it difficult to precisely ascertain the right value of $\hat{N}$. One can take the average of all three log-linear models, but then, why not include all models with dependencies. Fortunately, the use of MCMC simulations takes care of this issue.

Figure 3.5: Histogram of $\hat{N}$



Note: *The straight vertical line represents the 50th percentile and has a value of $\hat{N} = 4938$. The dotted vertical lines represent the 2.5th and 97.5th percentiles, respectively.*

Figure 3.5 shows a histogram distribution of all possible simulated samples from the MCMC simulation process. The vertical line depicts the 50th percentile and has a value of $\hat{N}_{MCMC} = 4938$, which is larger than all the three top specifications of the Chao log-linear model. The confidence intervals at 2.5% and 97.5% are depicted by the dotted vertical lines and have the values of 3635 and 7829 respectively. Therefore any value falling between these intervals has a 95% probability of estimating the true value of N.

Based on the estimates of the event abundance from both the preferred Chao log-linear models with dependencies as well as the MCMC estimate, the average probability of capture by each source is calculated. These probabilities take into account the heterogeneous nature of the sample. Table 3.7 shows the probabilities of reporting events by each source. The first three columns depict the data from MCMC estimates, which act as the baseline detection (capturing) rate for determining the extent of the selection bias.

Generally, the probability of detection of violent events for all media sources is very low. The rates retain a similar trend with the number of observations captured by each media source. The probabilities range from 6.1% [3.7%, 8.6%] detection rate for AFP (highest) to 0.5% [0.3%, 0.8%] detection rate for SANA (lowest). These values of the selection biases for each source are at significance levels of 5%. This implies that it is safe to conclude that SANA's detection rate is less than 1% of the actual violent events that took place within the selected time period. The rest of the media outlets have a detection rate of less than 10% including the confidence intervals. These capture rates reflect the magnitude of the selection bias prevalent in media reports on conflict. The last three columns of table 3.7 show the probabilities of detection for the preferred Chao models. Again, there are no large significance differences, as all detection values fall within the confidence intervals of the MCMC estimate.

In summary, the use of capture recapture models to estimate the true value of the observed and unobserved number of violence events in the specified period reveals the insufficiency in the use of media reports for coding conflict event data. These results are robust to various models, both using log-linear and Markov chain Monte Carlo estimations, and are significant at the 5% level. Moreover, the best performing models take into consideration the underlying heterogeneity between sources. This implies that the probability of capturing events vary from one source to the other. The heterogeneous

Table 3.7: Probabilities of the reporting selection bias

	$\hat{N}_{MCMC}$	CI(2.5%)	CI(97.5%)	$\hat{N}_{Chao(1)}$	$\hat{N}_{Chao(2)}$	$\hat{N}_{Chao(3)}$
BBCM	0.047	0.063	0.029	0.056	0.049	0.047
Al-Arabiya	0.055	0.076	0.056	0.067	0.059	0.034
AFP	0.061	0.086	0.037	0.075	0.065	0.062
AP	0.018	0.026	0.011	0.023	0.019	0.018
SANA	0.005	0.008	0.003	0.006	0.006	0.005

Note: *$\hat{N}_{MCMC}$=4938; 95% CI $\in$ [3635, 7829]; $\hat{N}_{Chao(1)}$=4085; $\hat{N}_{Chao(1)}$=4666; $\hat{N}_{Chao(3)}$=4880;*

nature of the data highlights the deliberate non-random selection of news reported in each source. Hence, the use of conflict data without accounting for possible selection biases is likely to produce biased results. In the next subsection the geographical aspects of the selection bias between the five sources is measured to determine the spatial consistency of media-based event data.

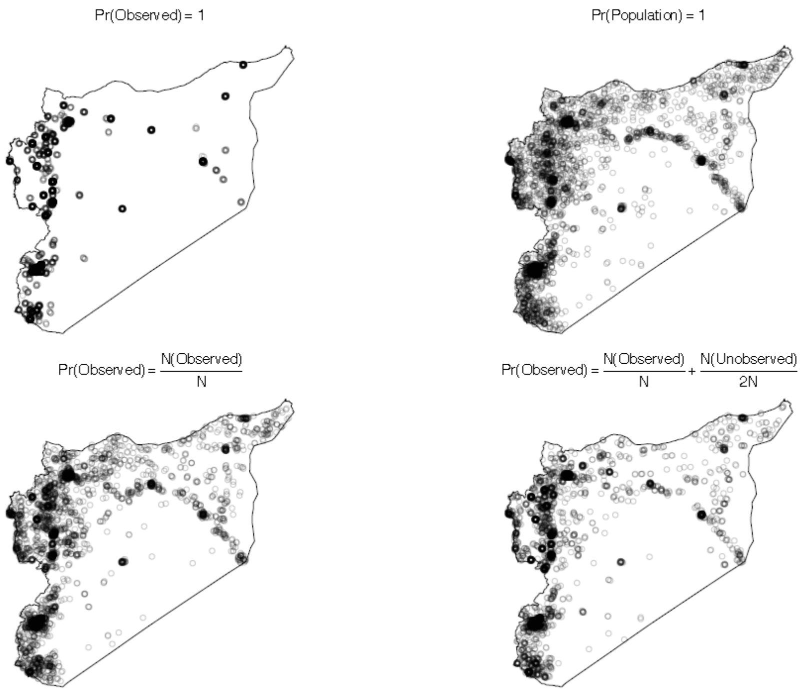
3.6.2 Spatial Intensity of the Selection Bias

As explained in detail in section 3.5, I simulate the spatial distribution of the unobserved violent events using the estimated total number of events $\hat{N}_{MCMC}$ from the capture recapture analysis. This allows comparing the spatial distribution of each media source to the likely true spread of violence in the selected time period.

Figure 3.6 provides a graphical example on how these processes are generated. The upper left part of the figure shows the simulated number of events $\hat{N}_{MCMC}$ based only on the coordinates from spatial distribution of the observed data, while the upper right figure shows the same number of observations based only on the coordinates from the spatial distribution of population densities of Syria in 2011.[18] These two processes showcase

[18] *The spatial coordinates of the population data are obtained from GAR 2015 exposure dataset (De Bono and Chatenoux, 2014).*

Figure 3.6: Inhomogeneous simulations of violent events based on various spatial intensities



Note: *For all the following figures, the Pr(Observed) + Pr(Population) = 1. Where $N(Observed)$ equals the total number of observed violent events in the dataset, and N is the abundance estimate of the total number of violent events. The probability shares determine the number of points from N, which are generated based on the underlying spatial distributions of the observed data and the population data respectively.*

the possible boundary distributions of violence. Hence, any combination of these two processes is likely to represent the spatial spread of the unobserved violent events better.

For the sake of demonstration, the lower left part of the figure keeps the distribution of the exact number of events from the observed data (856), and generates the coordinates of the unobserved events from the population data. The lower right figure, alternatively, generates the coordinates of unobserved events by applying equal probabilities between the distribution of the observed events and the population densities.

Finding the exact distribution of the unobserved data is not an easy task, yet checking the boundary intervals on various plausible distributions can shed light on the actual location-specific selection bias of each media source. To date, there is no research that approximates such distributions, or provide guidance on how to find the most optimal combination between observed data and fitted values of its covariates. The most suitable approach would be fitting a generalized linear model with a poisson distribution of observed data using the population densities, and then extracting the R^2_{Pseudo} based on the McFadden's computations. R^2_{Pseudo} is equal to $1 - \frac{D}{D_0}$, where D is the deviance of fitted model, and D_0 is the null deviance. Then the probability shares that determine the final spatial distribution of the data are based on the value of R^2_{Pseudo} and are calculated as follows:

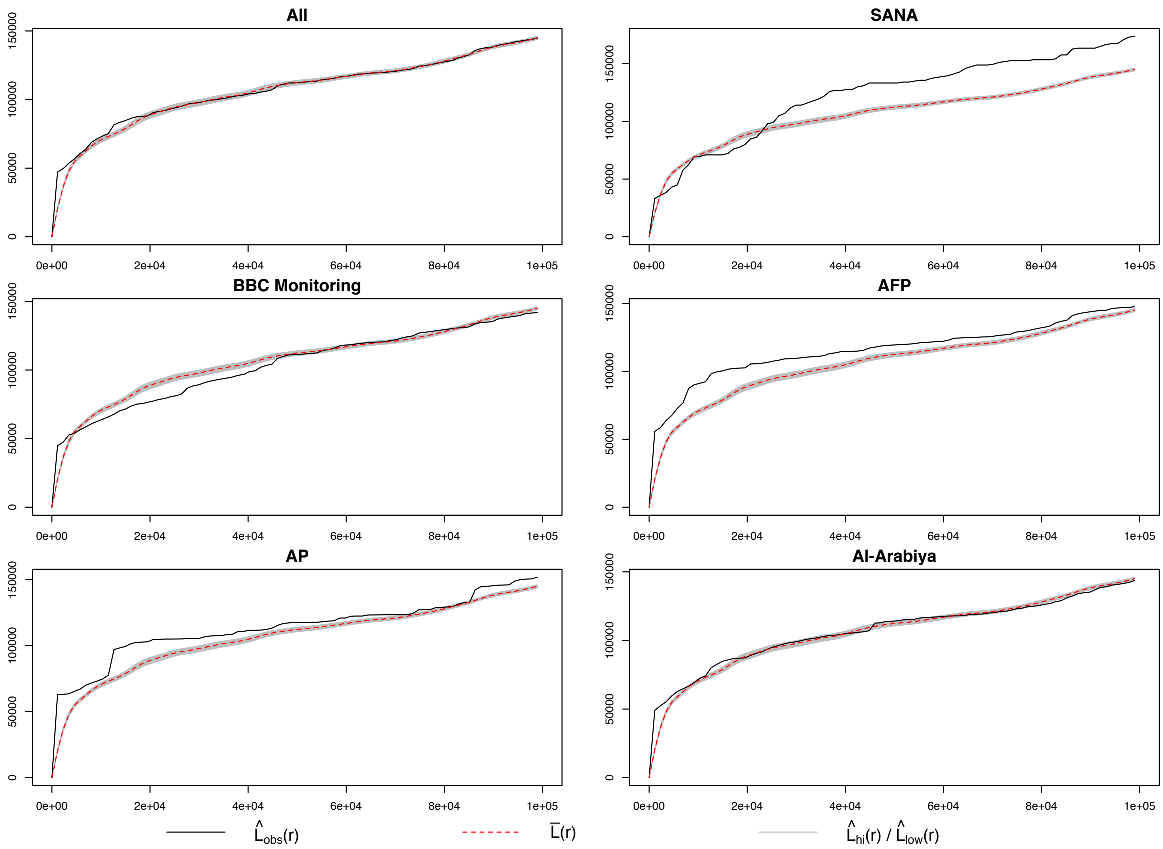
$$Pr(observed, population) = (1 - R^2_{Pseudo} + \frac{N(observed)}{2\hat{N}}, R^2_{Pseudo} - \frac{N(observed)}{2\hat{N}})$$

With an R^2_{Pseudo} equal to 0.31, then the probabilities determining the underlying distribution of violence is approximately to $\lambda = (0.79, 0.21)$. This implies that the 79% of locations of the simulated poisson distribution are drawn from coordinates of the observed data, while 21% are drawn from the population density estimates.

Each of these inhomogeneous distributions is used as part of the intensity parameter of the Ripley's L curve. I run 39 simulations to generate envelopes based on the probability distributions. The original spatial distribution of violence from each source is then compared to these envelopes. Any upward deviation from the envelopes implies that there is a positive bias (clustering of the data points), while a downward deviations implies a negative bias (repulsion of the data points).

As a comparison group, figure 3.7 shows the $\hat{L}$ curves, for which its envelopes are generated based on the intensity function of the spatial distributions from the observed data. The x-axis depicts the distance r in meters, and the y-axis depicts $\hat{L}(r)$. The envelopes around the dashed line shows the distribution of the intensity function based on the

Figure 3.7: L-Ripley envelopes with intesnity based on observed data



Note: *The x-axis depicts levels of r in meters, and the y-axis depicts the values of $\hat{L}$. $\hat{L}_{obs}(r)$ represents the L curve based on the observed data for each source. $\bar{L}(r)$ is the median (from the simulations) L curve of the underlying intensity function of the observed data, and $\hat{L}_{hi}(r)$ and $\hat{L}_{low}(r)$ represent the highest and lowest L curve (envelopes), respectively.*

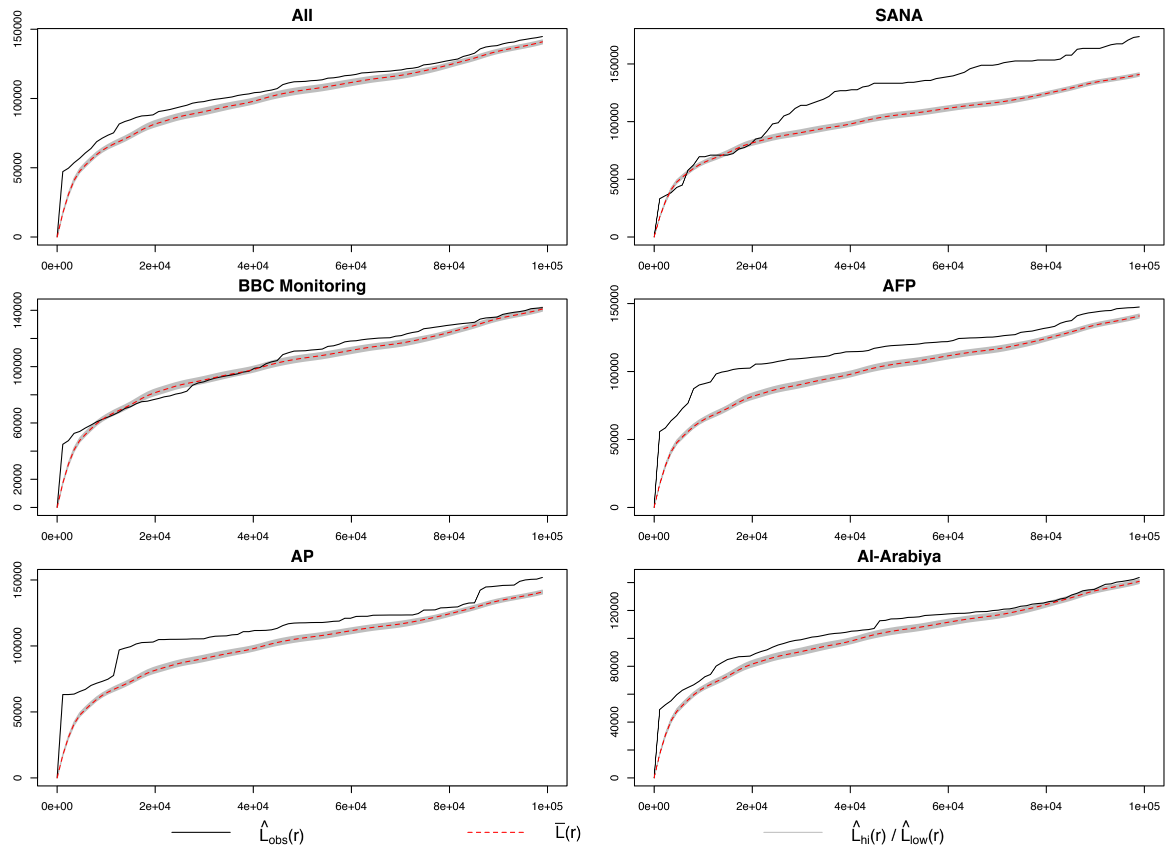
simulations. The black line, however, shows the distribution of the observed event data.

First, using all the data points in the dataset, the black line $\hat{L}(r)$ in the top left part of figure 3.7 falls inside the envelopes for most distances r. The notable deviation is found at distance values of $r = 0$. At this level, the observed data is clustered more than its expected it to be. This deviation is mainly driven from the fact that only the centroid coordinates of a city or town are recoded for most locations. Hence such a leap should not post a concern on the behaviour of the simulated samples. Second, AFP and AP exhibit a positive bias for all values of r, implying that the spatial distribution of violence events from these two sources are clustered. This clustering response is seen as well from the grid cell representation of violence in figure 3.4. For instance, in the case of AFP, the largest share of reported violence events were in the vicinity of Damascus. The $\hat{L}(r)$ from BBC Monitoring and SANA, on the other hand, reveal at lower values of r after the spike at $r = 0$ a repulsing behaviour. This dispersion implies that these two media source report less events at these distances than its expected. However, at larger distances, the curve of BBC Monitoring moves insides the envelopes of the intensity function distribution. In contrast, SANA falls back into the clustering trend at similar distances. Such behaviour from $\hat{L}(r)$ depicting data from SANA is expected given the low number of reported observations - which in turn effects the detection of events at various levels of r. Lastly, the spatial spread of violence events reported by Al-Arabiya behaves consistently with the underlying distribution of the simulated data. This is true for most most lower and higher values of r. Hence based only on the spatial distribution of violence from the observed data, Al-Arabiya has no significant spatial bias its reporting.

Yet how do these depictions change when generating the envelopes of the underlying function based on the preferred probability distribution of λ. Figure 3.8 shows the similar curves of $\hat{L}(r)$ as before, where the only difference now is the enveloped curve. After the inclusion of the share of population densities into the underlying function, the envelopes shift downwards. This implies that the spatial distribution of the observed data is positively biased. This trend is mostly evident in the top left part of the figure, where all the observed events are included. The clustering effect for the observed dataset is prominent for all values of r.

The most notable changes from the reference analysis are as follows: The violence events reported by Al-Arabiya are more clustered than initially predicted, particularly for lower values of r. In contrast to Al-Arabiya, the events reported by BBC Monitoring now exhibit no deviation from the expected spatial distribution of the violent events. For example, if a researcher would use only the distribution of the observed data (as one would be expected to do), then Al-Arabiya seems like the right choice in delivering

Figure 3.8: L-Ripley envelopes with intesnity based on λ



Note: *The x-axis depicts levels of r in meters, and the y-axis depicts the values of $\hat{L}$. $\hat{L}_{obs}(r)$ represents the L curve based on λ. $\bar{L}(r)$ is the median (from the simulations) L curve of the underlying intensity function of the observed data, and $\hat{L}_{hi}(r)$ and $\hat{L}_{low}(r)$ represent the highest and lowest L curve (envelopes), respectively.*

unbiased spatial inferences. However, in reality, the results would be upwardly bias as there is an over-prediction of violence occurrence at these locations. The researcher would be better off using the event data from BBC Monitoring. On average, all media sources exhibit area-specific biases at most values of r. This reconfirms the concerns regarding the prevalence of selection bias in using secondary sources for generating conflict event data.

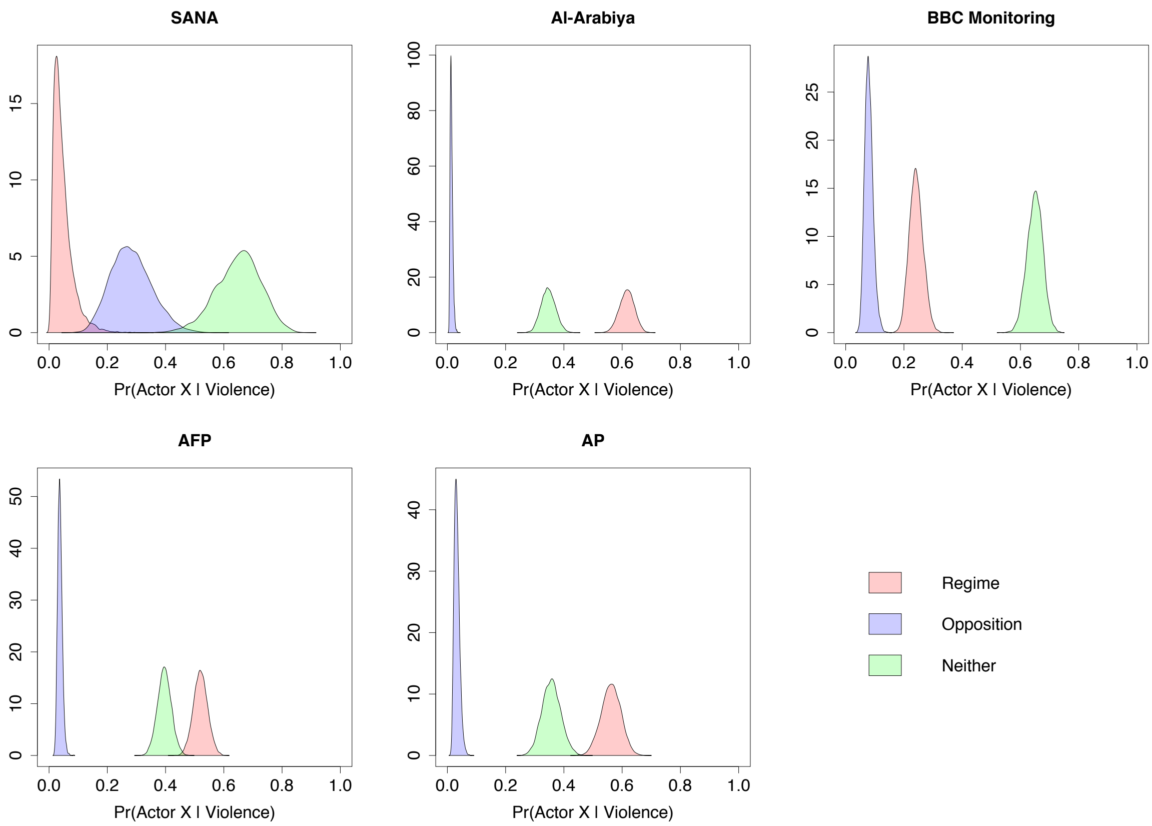
3.6.3 Actor-based Veracity Bias

For the purpose of this study, the veracity bias is only analysed on the reported armed actor(s) by each source. The bias is estimated as the joint predicted probabilities of reporting either Regime or Opposition. These predicted probabilities are based on the propensities from the bivariate logit model. The results from the analysis here are two-fold. The first estimation shows the expected values of these probabilities based on 10,000 simulations to examine the propensity of a media source to report either regime, opposition, both, or none, conditional on reporting any violent event. The second estimation shows the same propensities, but conditional on reporting violence against civilians. This latter distinction is important to capture the bias conditional on the type of violence perpetrated. If there are actor-specific description biases in the data, then it is most likely to be evident when violence is indiscriminate against civilians (Zhukov and Baum, 2016).

The density plots from the expected values of the predictions of the first estimation in figure 3.9 reveal a predictable trend. The propensity of SANA in reporting violence perpetrated by the government is relatively low (95% CI $\in$ [0.01, 0.14] compared to the violence perpetrated by the opposition (95% CI $\in$ [0.16, 0.43]). Yet the highest proportion of violence reported by SANA is, surprisingly, to neither actors. This effect is robust even for the inclusion of armed actors defined as "terrorists" as opposition forces. The density plots from Al-Arabiya, AFP, and the AP reveal close similarities. The propensity of reporting only opposition forces as the actor perpetrating violence is very low (close to zero). For example Al-Arabiya 95% CI $\in$ [0.01, 0.02]. Moreover, violence is more likely to be associated with regime or government forces. BBC Monitoring, on the other hand, is most likely to dissociate both actors from their reported violence events. The regime nevertheless is more likely to be mentioned as the perpetrator than the opposition forces.

The differences after controlling for the type of violence, in figure 3.10, reconfirms the

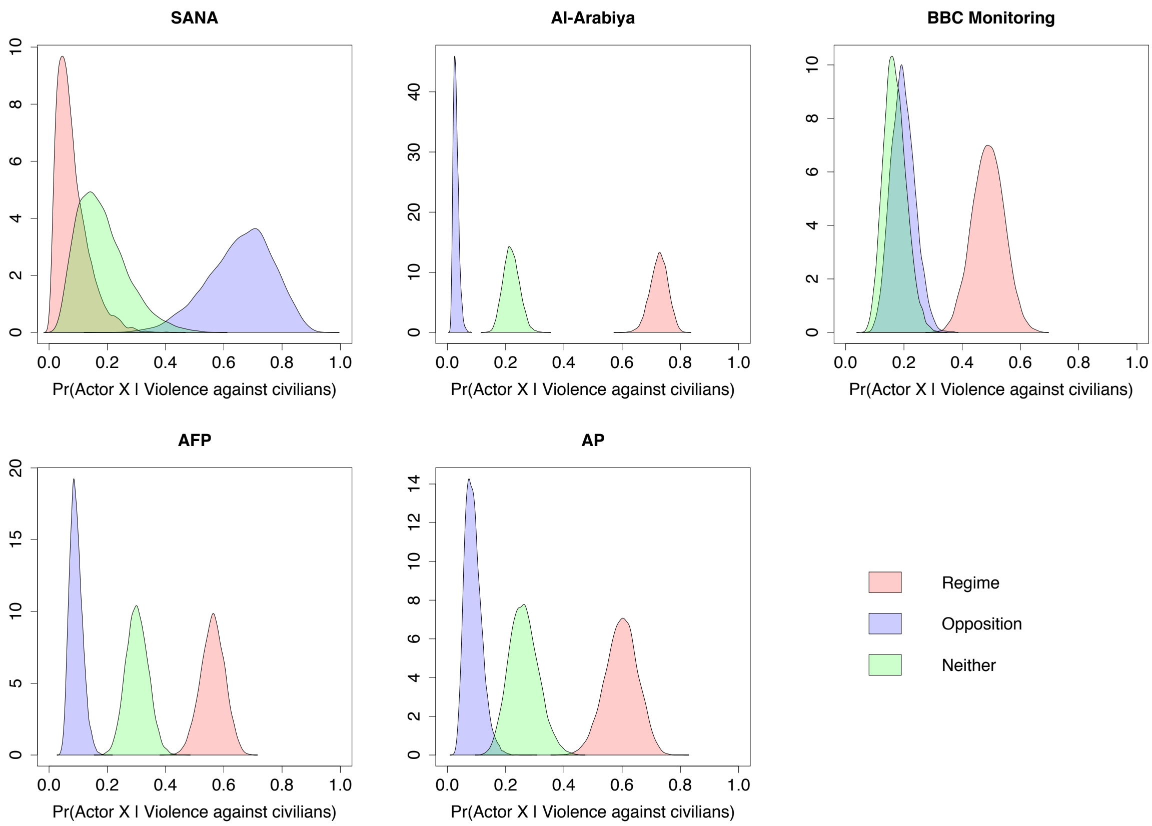
Figure 3.9: Propensities of reporting an armed actor conditional on reporting any type of violence



Note: *The densities of category "Both" is excluded in the figures to facilitate the presentation of the results. The sum of all predicted probabilities is equal to one. The y-axis values are not uniform across all sources, as the focus is to show differences within each source.*

results from the initial analysis. Here the density plots show the propensity of reporting a certain actor, both, or none, conditional that the violence is targeted against civilians. Despite the presence of no systematic differences from the initial estimations, three notable findings are worth mentioning. Firstly, in SANA, the propensity to associate the opposition with violence perpetrated against civilians is larger and more evident. Secondly, in Al-Arabiya, the propensity to associate the regime with violence targeted against civilians is more pronounced. Lastly, the propensity for BBC Monitoring to associate violence against civilians to neither groups drops significantly, and the gap in the expected values of Regime and Opposition is narrowed.

Figure 3.10: Propensities of reporting an armed actor conditional on reporting violence against civilians



Note: *The densities of category "Both" is excluded in the figures to facilitate the presentation of the results. The sum of all predicted probabilities is equal to one. The y-axis values are not uniform across all sources, as the focus is to show differences within each source.*

3.6.4 Interpretation of the Results

The results stemming out of the previous analyses recapitulate the deficiencies in relying on secondary media sources for the collection of conflict event data. This is partly driven by the fact that media sources fail to detect a large number of violent events that take place in a given war - at least in a protracted scenario as that of Syria. The detection rate of violence for most of the prominent news agencies is well below 10% at a significance level of 5%. This raises substantiated concerns for both researchers interested in studying conflict, and practitioners in need of reliable evidence to act.

Continental traditional media sources, such as the Associated Press and Agence France Press, produce very similar reporting trends. Both news agencies report violence in limited geographic areas (mostly urban), and are less likely to attribute specific types of violence and actors to these events. Moreover, about half of the reports from AP do not provide detailed locations, hindering a high precision geo-coding of violence. The carefulness in providing few details in their reporting of conflict is justified, as it underlines the credibility of their news. However, the absence of these details are a challenge for conflict event datasets.

BBC Monitoring did not report a large number of events as initially anticipated. The reporting technique used by the BBC Monitoring increases the capturing of local newspapers and broadcasts. However, this advantage of BBC Monitoring was not sufficient to detect larger number of events *vis-á-vis* other international sources. Nevertheless, as demonstrated earlier, it generated a better spatial coverage of the violent events in comparison with the rest of the international media - which is shown to be a recurrent issue (Kalyvas, 2004).

The advantages in including local and regional sources, such as Al-Arabiya and SANA, are not fully discernible. SANA reported far more fewer violent events than the rest of the included sources. Most of these events favoured the regime in Syria. In contrast to SANA, Al-Arabiya reported a larger number of events, yet it favoured opposition forces, associating most of the violence committed against civilians to the government forces. Prior work found that regional and national news sources exhibit a preference towards covering local events (Woolley, 2000). However, in the case of Syria, this reasoning does not seem to hold. The absence of independent news sources in Syria outside the government control (at least in the early years of the conflict) limits the availability of information, and subsequently a sound analysis.

The potential advantage of local sources in detecting more events is outweighed by its quality. As Zhukov and Baum (2016) consistently found, and as reiterated again in this chapter, local and regional newspapers significantly alter the veracity of their news (e.g., involved armed actors) contingent on their political inclinations. Hence, the addition of local sources to the bank of information used in the event data generation might increase event prevalence, but is likely to intensify the bias.

One would not expect that every media source broadcasts all violent events taking place, particularly in high intensity protracted conflicts such as in Syria. However, the results in the first part of this chapter show that, even with low detection rates, the heterogeneity of reporting events varies significantly between one media source and the other. Hence, dependent on the pool of sources used by a given event dataset, there is a high likelihood that the absence of some events is not driven by pure chance - rather from the deliberate exclusion of these events. Moreover, from an analytical perspective, deviations from the absolute number of events would not pose a risk on the analysis outcomes as long as the distribution of the events spatially resembles the reality. Nonetheless, as shown from the results of this chapter, selection biases are most likely to also be geographically biased. The source-based spatial differences hinder any plausibility in applying statistical corrections to improve quantitative analysis, as argued for by Weidmann (2016). This problem, thus, can present researchers with additional analytical challenges.

Both of the renowned and widely used conflict event datasets, ACLED and UCDP-GED, use secondary media sources in their coding processes. According to the UCDP-GED, their list of media sources varies from region to region, yet it always includes BBC Monitoring and one of the major news sources (Reuters, AFP, Xinhua, EFE). The coded events are scrutinized and judged manually by the coders and/or supervisors for possible biases in reporting. ACLED gathers its material using a number of sources, but primarily also through local, regional, and continental media. In ACLED, each event is coded from at least one source, and if more than two sources capture the same event, then the most thorough report is used. None of these datasets, however, offer a thorough assessment of the possible biases in their source selection. The matching analysis and capture-recapture models, which have been used in other disciplines, and newly developed and augmented in this chapter for testing conflict events data, can offer a minimal methodological solution. Therefore, it is recommendable that conflict event data collectors and coders include as much events possible from their pool of source materials. By doing so, they are able to calculate detection rates to determine selection biases to provide, at the least, a better insight on the nature and quality of the data they produce.

Yet detecting these biases is just the first step in a demanding process of providing comprehensive and reliable data. Correcting for these biases, in addition to solving issues associated with the use of secondary media sources (e.g., in relying heavily on human coders), is a challenge in itself. Due to the lack of rigorous validation and verification methods for such data, researchers face severe limitations with its analysis. The generation of quantitative viable weights (other controlling for event source in regressions) that account for these biases would be a step in the right direction. Yet given the difficulty in obtaining information on the primary sources, the development and availability of weights are meagre. In the next part of this chapter, based on a successful pilot, I propose an alternative method to the collection of conflict data. By working directly with primary sources on the ground, reporting fatigue and media biases can be completely eliminated. The exclusion of media as sources of information on warfare could be the golden ticket for conflict and peace research.

3.7 Crowdseeding: A Novel Approach for Collecting Conflict Event Data

Alternative methods to secondary media sources in collecting geo-coded conflict event data are scare and remain well underdeveloped. Despite the promised potential in using machine coding to go through large bulks of sources, which does reduce time and costs *vis-á-vis* human coding, many challenges remain. This is particularly evident, given that the selection and description bias is entrenched in the sources themselves.

During the last two years, a team from the London School of Economics (LSE) and the International Development and Security Center (ISDC) including myself, initiated a novel methodology for the collection of conflict event data. The technique builds on the pilot work by Van der Windt and Humphreys (2016), which is named crowdseeding, yet but taking a more rigorous and structured approach in the generation of event data. The remainder of this chapter will introduce the methodology and its codebook briefly, discuss its advantages in comparison to traditional methods, and finally propose a way forward on how such technique can be upscaled to generate reliable and consistent conflict event data.[19]

3.7.1 Introducing Crowdseeding

Crowdseeding works directly with primary sources on the ground during a conflict to obtain day-to-day event data on violence and other political and peace actions taking place locally. These primary sources (or seeds) are formed from a well-trained selected team of civil society activists residing and distributed in a given country under conflict. In the current project, we have trained eleven reporters, whom are spread across multiple regions inside Syria. They were requested to report on incidents that occurred in their vicinity, and enter the data daily directly to the research team. In order to provide a convenient, simple, and secure way to obtain the data, a designated online platform was developed for this purpose.

The reporters have access only to the entry sheets in the platform. These entry sheets are developed to ensure speedy, yet reliable entry of the data. They are also structured in line with the codebook developed for the project. Hence, the reporters fill in the sheets by only ticking the correct boxes. The back-end program of the platform then automatically transforms these entries into quantitative data, which with minor cleaning efforts are readily available for analysis.

To summarize briefly, the data collection using crowdseeding proved demanding in its

[19] *Unfortunately, given that the data collection is still under way during the writing of this dissertation, the descriptive presentation and use of the data is beyond the scope of this work. A published report on the dataset will be released around mid 2017. The report will be made available upon request.*

early stages. Given its novelty, and the nature of the war in Syria, challenges in regards to providing security and building a trustworthy team (trust in both directions) were present. However, after overcoming these challenges, the data collection went smoothly for period of nine months. During these months, we have gathered over 3,000 precisely geo-coded events on the conflict in Syria. The type of these events are presented briefly in the next section.

3.7.2 Type of Events

There are four type of events in the crowdseeding project on Syria: violent events, peace events, kidnapping, and looting. We narrowed down the choice of events to these options as they are most easily observed. The most notable difference between these types of events and that of other datasets, is the geo-coding of peace. Peace events include a variety of information on peace agreements and negotiations undertaken at the local level. Below, I will only briefly discuss the first two event types: violence and peace.

Violent events involves all actions that employ the use of force by an armed actor (either through the use of manufactured or non-manufactured weapons) that results in death, bodily harm, destruction, or displacement. The definition of a violent event is closely related to that utilised by DISC, yet minor variations have been implemented from the shortcomings of the secondary sources data collection.

A violent event has three sub-types: (i) violence between two armed actors; (ii) violence by an armed actor specifically targeted against civilians, and (iii) violence without a clear target or one-sided violence. The differentiation between those three sub-types of violence rely on a number of factors. The categorisation partly depends on the outcome. For example, if it is known to the reporters that civilians were targeted intentionally, then they would clearly select option two, otherwise they would select option three - no clear target. Moreover, the third category includes violence that is one-sided. For example, in case of a shelling targeted at a certain area without clear knowledge on the target group. These include mainly heavy shelling and aerial attacks. Violence is only then categorized as a battle or combat, when at least two opposing armed actors are involved. One main disadvantage of such a flexible categorization, is that it prompts reporters to rely heavily on the third option. However, with its exclusion respondents are forced to make arbitrary choices in the absence of sufficient information.

The tools of violence use the same list as that of DISC, which mainly includes: aerial attacks, heavy arms and artillery, small and medium arms, stationary explosive devices, and suicide attacks, *etc.* The outcomes of violence include deaths or injuries (classified by target), destruction (classified by type of dwelling or infrastructure), displacement, and territory gain (additional information on the actor and coordinates of the gained territory are provided).

Peace events include locally conducted talks and agreements related to both violence and non-violence. A peace event can take place between two armed actors, or between an armed actor and a political or civil society actor. We identified seven types of interlinked peace events. These include: (i) Invitation to talk about an agreement; (ii) rejection of invitation to talk about agreement, (iii) talks about an agreement: (iv) break-off of talks; (v) agreement reached; (vi) violation of an agreement (terms of agreement violated but the agreement is still considered to be in place); and lastly, (vii) beak-down of an agreement. A proposed agreement can take one out of these four metrical options: temporarily or permanently, and involving violence (e.g., ceasefire) or non-violence (e.g., delivery of aid). In addition to the open-ended list of armed actors attached to violence events, a number of civil society, international and judicial organizations were added to peace events to provide a larger range of relevant actors.

Both peace and violent events are geo-coded at the highest precision levels. The platform included an embedded map, where reporters simply clicked on the approximate position of an event. The coordinates of the event are then automatically recorded. Moreover, we additionally asked the reporters to specify the precision level of these locations. If they chose for example "at the district level" of a city, then the researcher knows that this exact point does not represent the coordinates of it location directly, and hence corrections are implemented accordingly (e.g., by taking the centroid coordinates of the city district). Other generic information includes: the date, the sensitivity of the event entered, the degree of observation, and the trustworthiness in the source of information if the event is not observed directly. The benefits in the inclusion of the latter two questions are discussed in detail in the following sections.

3.7.3 Methodology and Approach

The most novel aspect of the crowdseeding data collection is its methodology. Crowdseeding data is a form of organized crowdsourcing of information, where only a handful of selected and well-trained individuals insert data to a designated platform. The group of correspondents were selected based on existing networks within the project participants, who worked closely on research projects on Syria. Each individual received an online training session on how to use the platform data entry sheet through the team coordinator. Moreover, the project was piloted for a period of three months to ensure that its technical and conceptual aspects were fully understood by the team. During this period, the correspondents reported back any issues related to classifying events and their attributes, as well as on any glitches in the platform.

The data entry form was developed such that it only shows new categories based on previously chosen answers. This simplified and sped up the work of the reporters without jeopardizing the quality of the data entered, as it minimized errors involving the entry of unwanted data. Both the date and time of submission were recorded automatically by the system, while the event dates were coded by the reporters. The advantage of having precise submission times is manifested in the generation of reliability weights, as it enables an exact calculation of the time difference between the entry of the event and its occurrence (see subsection 3.7.4 for a thorough explanation on the usefulness of these weights).

Each reporter was designated with her own login account to the platform, which was highly encrypted. In order to strengthen the anonymity, only designated IDs were given to each reporter. Any personal information of the reporters were not recorded in the platform. Moreover, the project was vetted by security experts at the LSE to ensure the safety of the reporters and the data are maintained at the highest standards. To date, there were no breaches or incidents involving theft of data or exposure of the identity of the correspondents.

Building trust between the researchers and the reporters was challenging at the start of the project. Reporters were primarily concerned about the nature of the project. These concerns were mitigated by reiterating that the data will be used only for research purposes, by building strong relationships through common contacts, and by designating a trusted coordinator to undertake all communication directly with the team.

The coordinator also reviewed the entries on a bi-monthly basis to monitor the activity of the reporters and check if there were any major problems and issue in the data entered. Moreover, a bonus system based on the frequency of activity was implemented. This bonus system was not based on the number of events entered, but rather on the frequency of submission. The system aimed to motivate reporters to enter data on more frequent basis, rather than retrospectively at a weekly basis. Some reporters did not have internet access in some times, and all discrepancies or lack of activities were followed up thoroughly by the team coordinator. This was particularly important when there were no incidents of violence taking place for longer periods of time.

The commitment and involvement of the reporters extended beyond the sheer monetary reward. There was a need to reveal information to the outside world on what is happening in Syria. For example, some reporters who had to flee their houses at certain periods informed us immediately. Others who did not have plenty of events taking place in their vicinity requested to expand their areas of coverage by including their own network of friends. Their committed involvement was especially evident during the cease-fire in March 2016. Reports on the violation of events were entered on a daily basis, describing precisely what happened in a very detailed manner.

3.7.4 Merits and Limitations of Crowdseeding

The data generated via crowdseeding resembles closely that of other conflict event datasets given that it is built on a codebook. Therefore event data coming out of this technique can be a viable alternative to media-based resources, as it will not affect the end results. The advantages of crowdseeding over other types of data collections effort are many-fold. Table 3.8 compares a short list of these advantages *vis-á-vis* crowd-sourcing and conventional methods.

First, crowdseeding is undertaken on a day-to-day basis, as reporters submit the events when they see them happen. Although media sources report events in short time intervals as well, the coding process is usually performed in a retrospective manner. Coders could face issues, especially if they need to reconfirm the details of some of the reported events.

Second, all data types submitted via crowdseeding are geo-coded at the highest precision levels. As mentioned earlier, reporters only supply information on the incidents in their vicinity. Moreover, the ability to use an embedded map to select the location helps in determining, with very minor errors, the exact coordinates of an event. 25% of events from news reports, as shown earlier in this chapter, either reported incidents at the governorate or district level. Moreover in geo-coding locations of cities and towns, I always had to choose the coordinates of their centroid location. This severely limits the type of spatial analyses that can be performed with such data.

Third, given that crowdseeding works directly with the primary sources, media bias does not pose challenges any more. This is also true for the elimination of reporting fatigue associated with media reports, particularly in protracted conflicts, such in Syria. However, I am completely aware that biases can also be prevalent in crowdseeding. The verification of the information provided outside the pool of reporters is difficult to achieve. Nevertheless, solutions to these issues can be embedded in the design of each project. Dependent on the project capacity and availability of financial resources, it is possible to work independently with more than one reporter in a given area, and cross-validate the

Table 3.8: Advantages and disadvantages of crowdseeding

	Crowdseeding	Crowdsourcing	Traditional
Non-retrospective	✓	✓	
Elimination of media bias	✓	✓	
Elimination of reporting fatigue	✓		
More detailed and micro	✓	✓	✓
High Location Precision	✓		
Codebook-oriented	✓		✓
Less costly		✓	
Less Risky		✓	✓

information they provide. This increases substantially the costs of the data collection. However, with time, each project initiative can then narrow down their team based on their experiences in setting a trustworthy team to provide high quality service. This problem can also be driven by the political inclinations of the reporters. Working with teams living across different areas of control can perhaps shed more evidence on such biases.

In SCAPE, most reporters resided in areas controlled by the opposition. Therefore, there has been initially tendencies to report more violations by the regime over the opposition. We aimed to hire a geographically and politically balanced team, and reduce coverage to areas controlled by extreme groups, such as ISIS. Yet the short project duration was disadvantageous in that regard, as building a strong team required consistent reassurances, especially in securing payments inside Syria. However, I believe that for similar projects with longer time frames, the researchers are able to form a good team of reporters. During the pilot phase, the internal monitoring system allowed us to double check the data entered, and inform the reporters in case of consistently biased data was provided. We then clarified to the reporter that any provision of false or biased information is actually counter-intuitive to its purposes. That is, in the presence of biased information, we are unable to draw a clear picture on what is happening in Syria, which subsequently hinders the provision of viable prescriptions for understanding the conflict. From the point of view of researchers, such validations and reassurance are hard to disentangle. Yet technically speaking, crowdseeding offers other tools to correct for plausible biases. This is done by providing the end user with additional control variables, which can be used for setting up weights for increasing the validity of the data.

In order to measure the accuracy of an event, a number of variables where included, which can be used to generate reliability event weights. These include, the proximity of the event, the timing of the submission, the source of information, and the trustworthiness of the source.

First the proximity of the event is calculated as the geographic difference between the location of a reporter and an event, determined through their respective coordinates. In order to minimize risks and ensure the safety of the reporters, their exact geo-spatial location through the IP addresses was not recorded in this project. However, the location can be captured through a proxy. During the deployment phase, we knew the area of operation in which each respondent was covering. Hence, we can use the centroid of the polygon of this area as a proxy measure of the reporter's location. Although this method can lead to minimal measurement errors with larger polygon areas, it still offers a close and precise approximation. Dependent on the security requirements of each project, if permitting, recording the exact coordinates of the reporters are of crucial value.

Second, we record both the time and date of an entry, alongside the date of the event being reported. These criteria were necessary in case reporters did not have internet access for longer periods, and there were events that needed to be entered retrospectively. In order to minimize the retrospective bias in such cases, we can calculate the temporal difference between the reporting date and the event date.

Third, we asked if the respondent witnessed (heard or saw) the event first hand, or knew about it from other sources. The other sources included the following two options:

"Heard from someone who witnessed the event personally" and *"Heard from someone who did not witness the event personally"*. In case the the reporters selected one of those two options, they were prompted to specify the trustworthiness of the source. The trustworthiness variable is based on a 5-point Likert scale ranging from *"Fully Trusted"* to *"Not at all"*.

Based on these variables, we are able to generate meaningful weights to account for the reliability of the reported events. Three possible examples of such weights are presented here, namely the trust weight, the geographic weight, and the temporal weight.

The trust weight depends on the source and the trustworthiness of this source, and is calculated as follows:

$$W(S,T) = \begin{cases} \frac{1}{T \times S} & \text{if } S \geq 1 \\ 1 & \text{otherwise} \end{cases}$$

where S is the source of the event, and $S \in \{0,1,2\}$ in this case. $S = 0$ if the source is the reporter, otherwise S depends on the degree of persons providing the information. If $S > 0$, then T measures the trustworthiness of the source, where $T = 1$ is *"Fully Trusted"* and $T = 5$ is *"Not at all"*. Hence the both T and S act as an inverted multiplier to the accuracy of an event.

The geographic weight depends on the location of the event and reporter, and is calculated as follows:

$$W(r,d) = \begin{cases} \frac{r-d}{r} & \text{if } d \leq r \\ 0 & \text{otherwise} \end{cases}$$

where r is a radius of the predefined buffer around the reporter's centroid coordinates; d is the distance between the centroid and the event. r can be chosen arbitrarily by the user to fit their needs taking into consideration the unit and level of analysis. The larger the distance between an event and the reporter, the lower the weight is given to an event. If $d > r$, then the distance between an event and the respondent is greater that the threshold radius. In this case, the event will be excluded from the analysis.

The temporal weight depends on the date of an event and the date of its submission, and is calculated as follows:

$$W(t_e, t_s) = \begin{cases} \frac{1}{t_s - t_e} & \text{if } t_s \neq t_e \\ 1 & \text{otherwise} \end{cases}$$

where t_e is the event date and t_s is the submission date. The larger the gap between t_e andt_s, the smaller the weight. The weight is equal to 1 when the difference is zero days. The inclusion of the submission time can also be beneficial. For example, the time interval of the day the event took place (e.g., Morning between 6:00 and 10:00) can produce a better temporal precision. This then allows the generation of more reliable weights.

The inclusion of any combination of these weights in the analysis can be effective in reducing concerns on the reliability and accuracy of the events. Hence, the advantages of crowdseeding outperform that of media-based resources, not just by eliminating the media and reporting fatigue, but also through providing quantitative corrections to any possible biases that are absent in traditional datasets.

3.8 Conclusion

Based on two mutually exclusive data collection initiatives in Syria, this chapter has highlighted the major drawbacks facing the future of conflict event data, namely information sources. Using the first dataset, which was built from various international and local news sources, I find a strong evidence in the prevalence of selection and description biases. First, capture recapture model are used to estimate the total number of events for a pre-specified period. All news media reported less than 10% of the total number of events. Second, the spatial heterogeneity in the selection biases is shown to be consistent in most of the sources. Third, local and regional sources favoured some armed actor over others. These inconsistencies shed light on the weakness of media-based event datasets in providing reliable and accurate geo-coded information on violence.

Crowdseeding, on the other hand, eliminates these biases and provide more robust and geographically precise information on violence at the micro-level. Moreover, this approach allows the geo-coding of local peace events, which remains scarce to date. The use of primary sources that collaborate personally with researchers is an important step forward for the betterment of conflict event data. Crowdseeding is still in its infancy, yet it has capacity for further advancement. The future developments should address concerns regarding the security of the reporters, as well as generating viable techniques for validating the data. This is especially important in high intensity conflict zones like Syria. However, in small and geographically limited conflicts, this tool can be very useful.

Chapter 4

Predicting Temporal and Spatial Diffusion of Violent Armed Conflict

4.1 Introduction

Predicting both the onset and intensity of violence in conflicts is gaining substantial attention in the fields of social and political science, where a handful of studies attempts to predict future violence both at national and sub-national levels (Balcells et al., 2016; Blair et al., 2016; Goldstone et al., 2010; Weidmann and Ward, 2010; Zammit-Mangion et al., 2012, among others). This outburst of studies generated important insights on the spatial determinants and drivers of violence, the dynamics of violence in civil wars, and the predictive capabilities of theoretically-based models of violence diffusion.

The use of predictions in the field of conflict analysis is an important undertaking. The sheer difficulty in conducting experimental and semi-experimental studies on armed conflict forced scholars to rely heavily on observational data. Therefore, in the absence of valid identification strategies, drawing consistent and unbiased inferences from observational data becomes a myriad challenge (Weidmann and Ward, 2010). Predictions can be key in revealing what is not yet observed based on existing data given that they reduce confirmation bias and prevent overfitting. Moreover, evaluating explanatory statistical models using forecasts improves our understanding on the causes of warfare, particularly with the continuing availability of out-of-sample data.

Early studies attempting to forecast future levels of warfare focused on generating yearly risk maps. This was initially performed by predicting the outbreak of intra-state and inter-state conflicts for a given country and year (Reed, 2000; Goldstone et al., 2010). The increasing availability of temporally and spatially disaggregated data and improved estimation techniques increased the precision and accuracy of conflict forecasts (Ward et al., 2013). Conflict event data allowed researchers to study the significance of a handful of geographic and socio-economic predictors, including accessibility (road, mountainous terrain), ethnic composition, population size, and wealth (Raleigh and Hegre, 2009; Zhukov, 2012). Two main variables however, are consistently found to be robust predictors of violence, namely time and space. For example, the incorporation of violence in neighbouring locations generates better forecasts of violence onset, highlighting important aspects on the spatial diffusion of violence (Weidmann and Ward, 2010).

Despite the advances in this scholarly work, the usability and applicability of available prediction models in informing policy makers and humanitarian agencies remain underdeveloped and unclear (Cederman and Weidmann, 2017). The generation of valid predictions of armed conflict continues to face a number of methodological challenges. The difficulty in selecting a consistently valid unit of analysis and a robust set of statistical methods which control for spatio-temporal dependence are two examples of these challenges. Such weaknesses become particularly evident when attempting to obtain accurate out-of-sample predictions of armed conflict.

Beyond these methodological shortcomings, there is still little evidence on how armed conflict escalation is influenced by variations in the intensity of violence in neighbouring regions. Moreover, no studies to date tested how the frequency of violent activities of state or non-state armed actors affect the escalation of violence. Based on these methodological and research gaps, the objective of this chapter is to address the following two questions (i) Are spatial and actor-specific drivers robust predictors of violence escalation? And (ii) which of the currently available prediction model designs and tools are valid for forecasting not just the onset of violence but also its intensity?

Using grid level units of analysis based on high precision spatial data on violence from Syria, I test the appropriateness and validity of including violence in nearby locations as a predictor of both violent conflict onset and escalation. In contrast to country and administrative district levels, uniformly sized grid cells are robust exogenous apolitical units of analysis. Furthermore, spatial panel econometric techniques with fixed effects are used to estimate the spatio-temporal determinants of violence escalation. The fixed effect model generates better estimates given the highly likely presence of correlated omitted variables, where each grid then serves as its own control. To my knowledge, both these considerations offer novel approaches in the prediction of armed conflict.

The results of this chapter are summarized as follows: First, using various specifications of grid sizes and neighbouring spatial coefficients, an increase in violence intensity in neighbouring grids is found to reduce violence escalation. These findings underscore the diffusion of violence intensity with time. However, the significant impact of incorporating violence intensity in neighbouring locations does not enhance the predictive powers of the models. Both in-sample and out-of-sample forecasts reveal that models using only temporal variations of violence outperform the spatio-temporal ones. Second, an increase in the violent activity of state actors in previous periods de-escalates violence, while an increase in the violent activity of non-state actors intensifies conflict. Third, only the use of fixed effects models that control for spatial dependence produce consistent estimates. This implies that the inclusion of time invariant covariates such as location accessibility introduce bias to the results. The last finding highlights further the need to enhance both the design and methods applied in our predictive models to advance the frontiers of conflict and peace research.

The rest of the chapter is structured as follows: Section 4.2 describes in detail the current state of the literature, highlighting the models and tools used in the analyses. Section 4.3 presents the dataset, and section 4.4 presents the empirical methodology. Section 4.5 presents the results on the determinants of conflict, the in-sample, out-of-sample forecasts of both conflict onset and escalation, and section 4.6 offers a number of robustness checks. The last section concludes this chapter.

4.2 Literature Review

In recent years, empirical research on the spatial determinants of violence in conflicts increased substantially, highlighting the importance of geography on the spread of violence (Rustad et al., 2011; Buhaug, 2010; Raleigh and Hegre, 2009; O'Loughlin and Witmer, 2011; Buhaug and Rød, 2006; Buhaug and Gates, 2002). Civil wars differ crucially from traditional interstate wars, such that fighting and acts of violence take place irregularly and locally between armed actors (Kalyvas, 2005). These irregularities in the onset of violence prompted researchers to understand better why does violence in civil wars take place at those specific locations. Theoretical and empirical studies identified a number of determinants that drive violence onset, and partially escalation, in civil conflict. These drivers can be categorized as follows: First, geographic-specific factors, such mountainous terrains and natural land cover (e.g., Buhaug et al., 2009). Second, socio-economic and demographic factors in these locations including, for example, ethnicity, population size and wealth (e.g., Fearon and Laitin, 2003; Hegre et al., 2009; Raleigh and Hegre, 2009). Third, the accessibility to these locations such as the proximity to borders and road networks (e.g., Zhukov, 2012). Fourth, climatic factors, such as temperature and precipitation anomalies (e.g., Hsiang et al., 2011).

Fitting models to predict the onset and intensity of violence using this comprehensive list of explanatory variables is not a trivial task. Statistically significant determinants of violent conflict are rarely found to have strong predictive power (Ward et al., 2010). Most of these variables are either time-invariant or slow-changing. Hence, for shorter periods of time (months, weeks, or days), variations in conflict intensity are less likely to be driven by (more difficult to be associated with) changes in these explanatory variables.

"Location, location, location" has been the eureka moment of Ward and Gleditsch (2002). In their paper, they apply an autologistic regression to account for spatial dependence and draw the out-of-sample predictions using MCMC simulations. Most importantly, the paper concludes that war outbreaks are mainly driven by other wars in their spatial proximity. Weidmann and Ward (2010), applying a similar approach, show that both spatial and temporal elements are vital for producing valid and robust predictions. The results underline the importance of incorporating violence in neighbouring locations to generate better forecasts of violence onset. However, to date, there are no studies that predict violence escalation with spatial dependence, or examine how does proximate violence impacts its escalation (contagion effect). These shortcomings come as no surprise, as the debate on the validity of the models and data used to generate predictions of violence is still inconclusive. Two interrelated aspects in this regard are highlighted for the purpose of this chapter. The unit of analysis and, consequently, the applied estimation techniques.

The choice of the unit of analysis affects both the availability of data and the accuracy in measuring the set of explanatory variables. For example, using smaller units of analysis diminishes the ability to use direct measures of wealth, or to provide adequate estimates of population densities. Therefore researchers face a clear dilemma in that regard: Using country-level or administrative-level units of analysis provide sufficiently accurate data, yet generates imprecise estimates. On the other hand, using smaller precise units of analysis generates good estimates, but data availability at these levels is scarce or insufficient.

Most studies analysing conflict onset at the national level were able to construct their dataset to incorporate most of these variables (Goldstone et al., 2010; Reed, 2000). Country level information on population size and GDP per capita are easily accessible and can be precisely integrated into the final dataset. This is also true for aggregated conflict event data on the district or sub-district levels. The downside, however, is that the use of spatial polygons defined by administrative or political boundaries disregards the importance of local attributes in driving violence, and relies on a strong assumption regarding the diffusion of violence. For instance, the distribution of violent conflict events is most likely not be homogeneous across a given country (Buhaug and Gleditsch, 2008). As shown in the previous chapter, even in the case of Syria, violence tends to cluster in specific geographic areas within the country. Hence, analysing the determinants of violence at lower units of analysis is vital to generate more accurate forecasts, particularly when using space as a predictor.

Alternative methods have been implemented to correct for such inconsistencies. Disaggregating the country area into uniform spatial pixels or grid cells (e.g., a square or hexagon grid) is gaining popularity in conflict research. The PRIO-GRID database is one of the pioneers on that front (Tollefsen et al., 2012). The advantage of using gridded data is demonstrated by the spatial design of the grid system. Pixels are fixed spatial units that comprise apolitical boundaries, and hence are truly exogenous to economic, political and social variables of interest. The disadvantage, however, is that there are yet no systematic procedures for choosing the correct grid size, which in turn impacts the outcomes of the analyses. Nevertheless, random allocations of grid sizes and consistent robustness checks are valid solutions in this regard - which will be discussed in detail in the methodology section.

Another approach involves using Poisson Point Models (PPMs). PPM techniques use the coordinate location of events themselves as the primary unit of analysis. This approach has been widely used in other disciplines, but only recently is being implemented by conflict researchers. Only few papers have applied this technique to forecast the onset and escalation of violence in civil wars. Zammit-Mangion et al. (2012) use dynamic Poisson models to estimate the predictive powers of the underlying processes in conflict, such as diffusion and heterogeneous escalation, and Schutte (2016) applies PPMs to positively cross-validate the predictive power of spatially dependent covariates usually found in literature. These models offer new insights on the possible approaches used for analysing and predicting violence in intrastate civil conflicts. They are particularly relevant as they mitigate challenges of falling into the trap of arbitrary choosing pre-specified spatial shapes. However, they disregard the temporal variability of conflict, as well as require large number of event data for the model to converge properly. Very recently, Blair et al. (2016) use household survey data in Liberia to draw predictions using a number of cross-validation approaches, such as the Least Absolute Shrinkage and Selection Operator (Lasso), random forests, and neural network analysis. They conclude that the use of the Lasso procedure produces superior results over the conventional methods (such as logit) for out-of-sample predictions. Yet the validity of these models in predicting escalation with longer time periods are still weak or untested to date.

In summary, the presented literature has implemented several methods to forecast the onset and diffusion of violence. Yet there remains notable gaps, particularly in using valid units of analysis, and applying robust models that account for temporal and spatial variation of violence relative to escalation or intensity. This chapter attmepts to close these gaps by applying fixed effect spatial panel analysis on grid cells (to correct for the prevalence of unobservables and spatial autocorrelation) to test whether space is a indeed a strong predictor of violence.

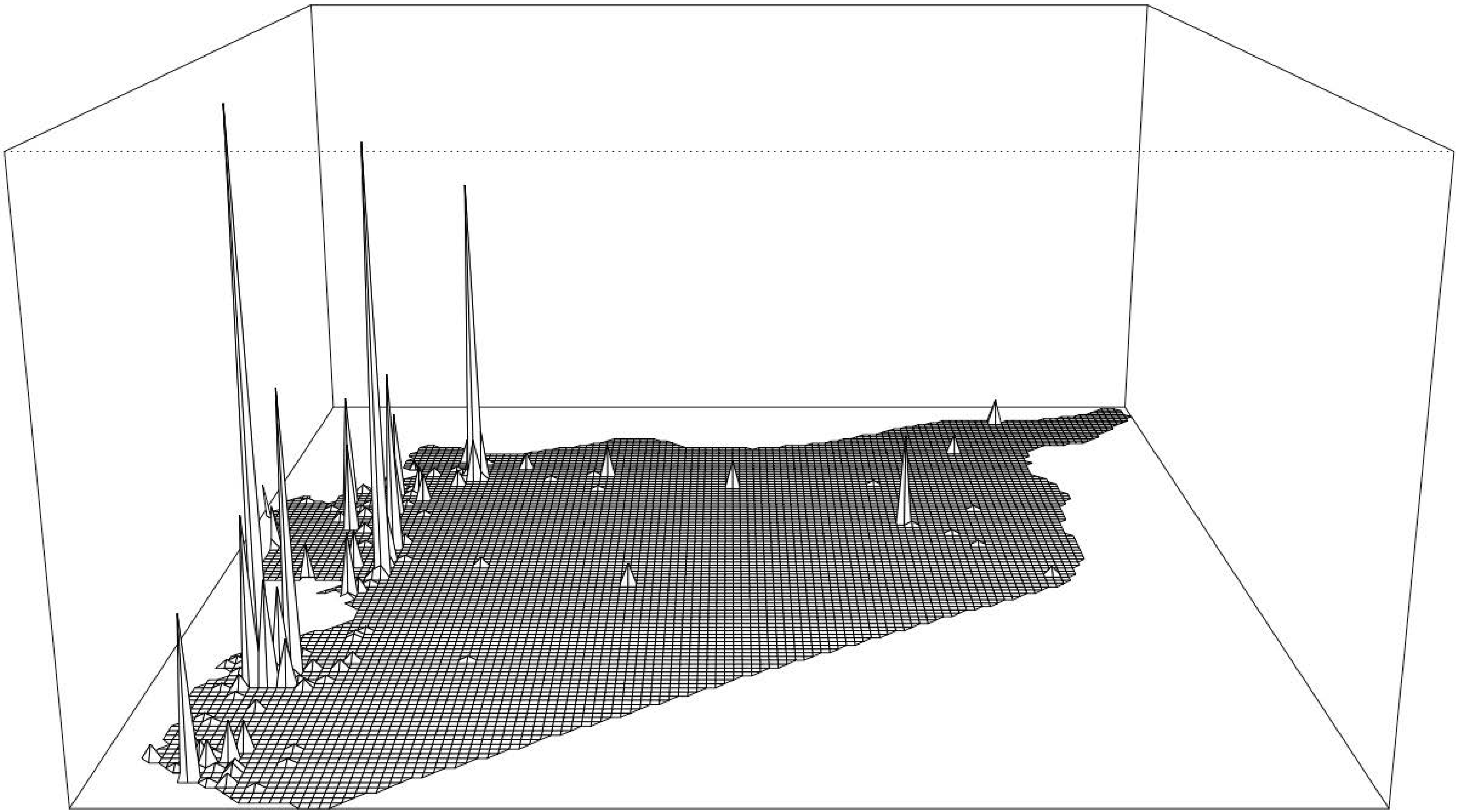
4.3 A Gridded Panel Dataset

The dataset used in the analysis is constructed using multiple sources. Violent events in Syria are extracted from DISC database. As shown earlier in chapter 3, the dataset does not include all violent (expected) events that took place within the sample time frame. However, the inclusion of all available secondary sources in the final dataset, although suboptimal, increases the representativeness of violent incidents. Events that are captured by more than one source are also given larger weights of occurrence. For the purpose of the current analysis, I am only interested in capturing the onset of violence (i.e., the geographic location of violence events in their respective periods) and its intensity (i.e, the number of events taking place in each period). Moreover, the involvement of armed actors in these events are also used as determinants of violence intensity. Hence, only three variables from the DISC dataset are of interest: the spatial coordinates of an event, the time period, and the actors involved.[20]

Figure 4.1 shows a three-dimensional spatial density representation of the point coordinates of all violent events from DISC. Again, the dataset includes only violent incidents with high geographic precision levels - where the coordinates can be clearly attributed to an event. These constitute about 76% of all violent events reported in the original dataset. As shown earlier in chapter 3, and evident from figure 4.1, most of the reported violence during the project's selected time frame is concentrated in the populated cities within Syria. This clustering effect is it to be expected in conflict event data (see for e.g., Buhaug and Gleditsch, 2005, 2008).

[20] *All actors are included from the matched events, and aggregated to the grid level, in order to account for possible description biases between the sources.*

Figure 4.1: Spatial density distributions of violence in Syria



The units of analysis are chosen to be the 25km sided grid cells, and the periods are the randomly selected 4-day segments (see table 3.1 for details on the study time frame). Hence, the final baseline dataset constitutes a balanced panel with 299 observations and 15 periods. The number of events falling within each grid cell, and for each period, are then aggregated to generate the variable *intensity*. On the other hand, the variable *onset* is constructed if at least one violence event occurs within grid x in period t. In other words, if a violent event occurred in a respective period and grid, then *conflict* equals to 1, and 0 otherwise. The number of actors involved for each x in $t-1$ is the simple sum of actors reported. The list of actors is aggregated into the following three categories: regime forces, opposition forces, and unknown forces.

In order to account for the possible spatial contagion of violence, two spatial neighbouring parameters are used, the rook and queen models. The names of these specifications follow that of the game of chess. The rook spatial lag, hence, only takes into account the horizontal and vertical neighbouring grids, while the queen lag, additionally, incorporates the neighbouring grid cells at the diagonals. Moreover, for each grid x, I include the following time-invariant explanatory variables: population size, wealth, access to roads, farmland, border areas, and variation in altitude.

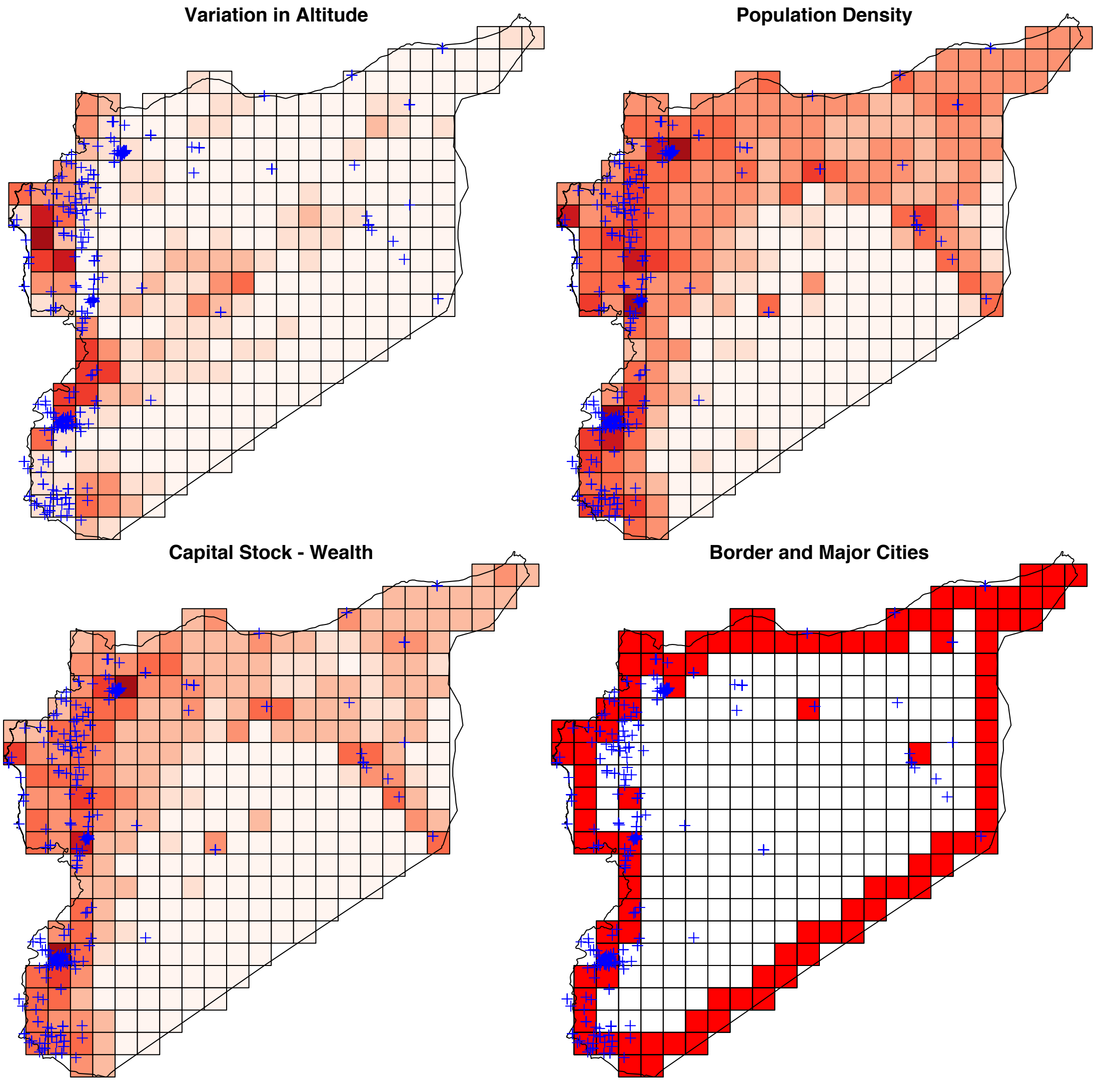
Population: Concentrations of civilian populations are strongly associated with the onset and intensity of violence at the local level (Raleigh and Hegre, 2009). The increase of violence in populated areas is theoretically shown to be related to a number of factors. These include, the dispersion of fighters within civilians (Salehyan and Gleditsch, 2006), needs for recruiting, and significant territory gain (Kalyvas et al., 2006). The data on population is adapted from the GAR 2015 exposure dataset for Syria (De Bono and Chatenoux, 2014). Proxies of population size for 5x5 grid cells are based on the 2011 population estimates. I aggregate the population figures based on their geographic information to the 25x25 grid level.

Wealth: Distributions of wealth is also found to be closely associated with outbreaks and escalation of violent conflict. Hegre et al. (2009) show that violence is more likely to cluster in wealthier regions, drawing on the possibility that rebels fight over areas with more resources. A spatial indicator of wealth cannot be disaggregated to the gird level in a straightforward manner. Hence, and in order to determine an appropriate value of *wealth* in each grid cell, I use proxies of capital stock from the GAR 2015 exposure dataset (De Bono and Chatenoux, 2014) as well. The downscaling process of wealth in the GAR exposure dataset consists in transferring the capital stock from the country level to the 5x5 grid cells. This downscaling approach is based on the regional variation of GDP, where for each cell a coefficient of variation from national values is calculated. Moreover, these values are then weighted by an estimated unitary value based on building topologies in each grid. The final value of wealth in each grid is then measured in thousands of dollars.

Accessibility: Mountainous terrain enables rebels to evade fighting, and hence prolong the periods of violence taking place in these regions, limiting the spread of violence to other areas (Fearon and Laitin, 2003). On the other hand, access to road networks play an important role in the spread and diffusion of violence as it facilitates the accessibility and movement of armed actors (Zhukov, 2012). Moreover, violence is more likely to take place close to the country borders. Borders have long been considered a vital route for insurgents to supply themselves with arms and recruits. Hence, three variables are included in the dataset to account for accessibility: variation in altitude, access to main roads, and the proximity to the country's borders. Data on the first two indicators are adapted from the DIVA-GIS open database (Hijmans et al., 2004). First, average altitude in a given geographic unit was widely used in prior studies as an explanatory variable (e.g., Weidmann and Ward, 2010). I strongly believe, however, that the roughness in mountainous terrain is better captured by looking at standard deviations from the mean altitude within each grid, rather than the mean itself. Second, only primary and secondary roads crossing through each grid cell are recorded. The variable takes a value of one if the grid has access to a road, and zero otherwise. Last, a border dummy variable is built such that it takes the value one if the grid cell is a border cell, either with respect to the neighbouring countries or the Mediterranean sea, and zero otherwise.

Land Use: A large proportion of geographic Syria is desertified, particularly in the eastern part of the country. Therefore, areas with fertile lands and farms are more likely to witness violence as actors aim to maximize their gains. Farmland cover is used an indicator of fertile lands. Similar to roads and border grids, a cell which contain fertile farmlands take the value of one, and zero otherwise. The data is also based on DIVA-GIS.

Figure 4.2: Gridded spatial distribution of the determinants of violence



Note: *Incidents of violence are indicated by the blue crosses. The coloured grids on the other hand, indicate for each sub figure: the standard deviation for the altitude, population, wealth, and and border areas. Legends are excluded for a better presentation.*

Figure 4.2 shows the spatial variations of some of the above-described indicators in more detail. They are mapped separately, and are spatially matched with violent events using the 25 km grid units. Descriptively, violence is likely to take place in populated and wealthier areas, close to the borders, and within major urban settlements. Nevertheless, mountainous regions seem to have less of an effect on the diffusion of violence.[21]

4.4 Methodology

For the purpose of this chapter, both spatial autologistic and multivariate methods are applied to predict the spatial and temporal diffusion of violence onset and intensity, respectively. The first approach is based on the model developed by Ward and Gleditsch (2002) and adjusted by Weidmann and Ward (2010). For the second approach (conflict intensity), however, I will mainly use econometric panel analysis with spatial error parameters. The advantage of this technique is demonstrated through its ability to measure spatial and temporal variation of violence intensity while accounting for grid fixed effects.

Before going into the details of the estimation approaches it is important to highlight an issue related to the use of grid cells. Given the tendency of violence incidences to cluster in particular locations over time, the use of econometric models on cell units has its advantages and drawbacks. The main analytical challenge is related to the size of the grid cell. The smaller the grid size, the higher the number of observation N, and hence the lower the standard errors are. Therefore, if the size of the unit of analysis is chosen too small, there is a higher likelihood that empirical coefficients are always statistically significant. Especially given the high number of zero observations (empty grids) in the data. On the other hand, choosing larger cells decreases the number of observations, but conversely diminishes the actual magnitude of the spatial inference in play. The geographic literature denotes this conundrum as the "Modifiable Areal Unit Problem"

[21] *It is important to keep in mind that most of the mountainous regions in Syria are areas which staunchly support regime forces, and to date have not witnessed violence incidence in the same intensity as in other parts of the country.*

(MAUP). The effects of MAUP are recently being considered in conflict studies (e.g., Linke et al., 2017). However holistic systematic fixes are still not properly developed. For the current analysis, I will rerun the analysis using various grid sizes, and check if the results are robust to changes in the choice of the grid. Moreover, and in order to account for the large number of grid with zero events, I undertake a famous Heckman approach in economics on self-selection (Heckman, 1979). In other words, I will test if the grids with zero violence are fundamentally different from the ones containing violence incidence.

4.4.1 Spatial Autologistic Approach - Conflict Onset

In order to examine the spatio-temporal determinants of conflict onset, it is vital to take into consideration the spatial dependence between observations, such that a conflict event $conflict = 1$ in neighbouring units at time t affects the likelihood that unit i experiences conflict. This approach is known as the spatial autologistic regression (Geyer and Thompson, 1992; Ward and Gleditsch, 2002). The autologistic model states the conditional probability P_{it} that $conflict_{it} = 1$, given values $conflict_{jt}$ at units $(j \neq i)$:

$$P_{it} = Pr(conflict_{it} = 1 | W_{conflict_{it}}) = \frac{e^{\alpha + \beta X_{it} + \gamma W_{conflict_{it}}}}{1 + e^{\alpha + \beta X_{it} + \gamma W_{confict_{it}}}} \tag{4.1}$$

where β is a vector of parameters for the explanatory variables, and γ is a parameter for the spatial lag of $conflict$. Hence, when $\gamma = 0$, equation 4.1 is reduced to a standard logistic model. Inversely, when $\beta = 0$, the expression is reduced to a pure autologistic model. W is a spatial weighted matrix which is based on the first order spatial lags of the grid cells. The use of grids allows the inclusion of two separate types of immediate neighbours: the rook and queen lags. Following the footsteps of Weidmann and Ward (2010), the second order spatial lags are excluded, as its spatial impact passes through the immediate neighbour, and hence it is captured by the first order lags.

Equation 4.1 is estimated using Monte Carlo random samples instead of the widely used pseudo-likelihood approach (Ward and Gleditsch, 2002). The MCMC estimation has been consistently shown to yield approximations closer to that of a full likelihood estimation compared to the pseudo-likelihood approach. A Gibbs sampler generates a set of simulated maps to find the values of the estimate that yield a sufficiently valid statistic.

4.4.2 Spatial Panel Maximum Likelihood Approach - Conflict Intensity

For the analysis of conflict intensity with spatial dependence, a general static spatial model is used that includes both temporal and spatial lags of the dependent variable and spatial autoregressive errors.

Let

$$y = \lambda(I_T \otimes W_N)y + X\beta + u \tag{4.2}$$

where y is an $NT \times 1$ vector of observations on the dependent variable, X is a $NT \times k$ matrix of the independent covariates, I_T an identity matrix, and W_N is the spatial weight matrix. λ is the spatial autoregressive coefficient of interest.

Both the random effect and fixed effect model of the general form are tested. However, I hypothesize that the fixed effect model is more consistent due the likely presence of correlated unobservables. Moreover, and as discussed by Elhorst (2003), the presence of the spatial lag introduces endogeneity, where the regressors are likely to be correlated with the error term. Corrections are based on Baltagi et al. (2007), such that the idiosyncratic errors are spatially autocorrelated, but the individual effects are not. Hence a fixed effect error model can be specified as follows (Millo and Piras, 2012):

$$y = (\iota_T \otimes I_N)\mu + X\beta + u \tag{4.3}$$

$$u = \rho(I_T \otimes W_N)u + \epsilon \tag{4.4}$$

where ρ is the spatial autocorrelation coefficient, ι_T a column vector of ones, and ϵ is a "well-behaved" error term. Equations (4.3) and (4.4) are then estimated by Maximum Likelihood (ML) to obtain the desired coefficients.

4.5 Main Results

4.5.1 Conflict Onset

The first results show the spatial and temporal determinants of conflict onset. I report the estimations of the Monte Carlo simulations with one degree spatial lag for both the rook and queen models. The results in table 4.1 compare these two models to the temporal baseline model, which uses loglinear estimations. The outcomes of the estimations are in line with the initial expectations. In all three specifications, the occurrence of conflict in periods $t-1$ and $t-2$ are both positive and significant at the 1% level. Hence, the temporal effect of conflict is a robust determinant of conflict onset at the specified grid level. Moreover, a larger population size increases the occurrence of conflict, while wealthier grids are less likely to drive conflict onset. The former outcome is in line with the previous results stemming from literature (e.g., Raleigh and Hegre, 2009). However, wealthy areas are found to be more prone to conflict (Hegre et al., 2009), which is in contrast the results coming out from these estimations. Moreover, geographic specific indicators such as the variation in altitude and access to primary roads are both positive and significant. This is also true for gird cells that are agriculturally fertile lands.

λ which denotes conflict occurrence in neighbouring grid cells is also found to be positive and significant at the 5% level. This is true for both the rook and the queen specification as shown in columns (2) and (3) of table 4.1. Hence conflict onset is also spatially driven by neighbouring violence even after controlling for temporal variations and grid-specific characteristics. However, the models in all three specifications do not largely vary from one another. In other words, the inclusion of the spatial lag of conflict does not seem to enhance notably the fit of the model. Therefore, predictions are undertaken to to test if the inclusion of the spatial lag enhances the performance of the model.

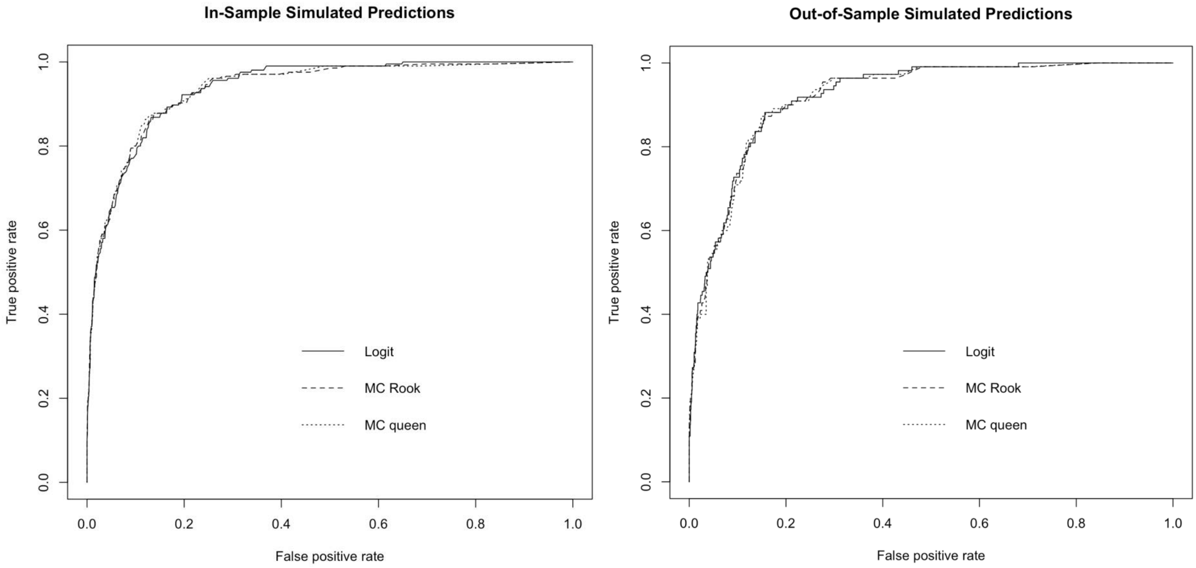
Table 4.1: Spatial determinants of conflict onset - baseline model

	(1)	(2)	(3)
Conflict_{t-1}	1.55^{***}	1.47^{***}	1.48^{***}
	(0.22)	(0.26)	(0.27)
Conflict_{t-2}	1.26^{***}	1.10^{***}	1.19^{***}
	(0.23)	(0.25)	(0.31)
Population	4.75^{***}	4.25^{***}	3.86^{***}
	(0.77)	(1.13)	(1.14)
Wealth	-0.20^{***}	-0.18^{***}	-0.16^{***}
	(0.03)	(0.05)	(0.05)
Roads	2.58^{***}	2.54^{***}	2.64^{***}
	(0.73)	(0.77)	(0.99)
Altitude (SD)	0.26^{**}	0.25^{*}	0.25
	(0.12)	(0.15)	(0.18)
Country Border	-0.03	-0.07	-0.13
	(0.20)	(0.20)	(0.24)
Farms	0.83^{***}	0.79^{***}	0.72^{***}
	(0.20)	(0.20)	(0.19)
λ		1.54^{**}	2.21^{**}
		(0.70)	(0.89)
Num. obs.	3887	3887	3887
Model	Logit	MC	MC
Spatial Lag	No	Rook	Queen

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard deviation in parantheses. The first spatial lag is denoted by λ, where specification (2) uses the rook model and (3) uses the queen model. Population in 100,000; Wealth in 1,000: Altitude in 100m; Roads, farms and country border are binary variables.

For all three models, two predictions are conducted to test the sensitivity and specificity of the models: the in-sample and out-of-sample predictions. Both predictions are based on MCMC simulations with 1,000 iterations and 100 burn-ins. The out-of-sample predictions are generated using periods three to nine as the training dataset, and testing them to periods ten through fifteen.

Figure 4.3: ROC - Receiver Operating Characteristic plots of conflict onset



***Note**: The Area Under the Curve (AUC) values for the in-sample simulated predictions for the temporal, rook spatio-temporal, and queen spatio-temporal specification are 0.937, 0.935, 0.936, respectively. As for the out-of-sample simulated predictions the AUC values are 0.937,0.919, 0.936*

Figure 4.3 shows the Receiver Operating Characteristics (ROC) plots for all three specifications, both for the in-sample (left) and out-of-sample predictions (right). ROC's are a convenient graphical presentation to convey the validity of the models. Their main advantage is that they do not rely on an arbitrary pre-specified threshold value to test the predictions. This is particularly important in conflict research as it is common that the number of observations including conflict onset are significantly small.

All specifications perform exceptionally well in predicting conflict onset. Both the in-sample and, the more stringent, out-of-sample predictions plots are strongly concave towards the true positives. All three curves are without doubt above the diagonal line. However, despite the excellent predictive capabilities of these models, there seems to be no notable differences between them. In other words, the incorporation of the spatial lagged variables of conflict onset, in both the rook and queen models, do not strengthen its predictive power. Even more, the Area Under Curve (AUC) estimations show that the temporal model outperforms both spatial models, while the queen specification slightly produces better predictions compared to the rook model. These initial results contrast the long-standing claim that the incorporation of the spatial occurrence of conflict helps explain future violent conflict onset better.

This analysis follows directly from that of Weidmann and Ward (2010), as it only looks at onset but not intensity. However, there are a number of methodological challenges that can affect the outcomes of the estimations, which are important to addresses. First, the use of political administrative divisions or pre-specified grid units can have an impact on the results. As mentioned earlier, recent work attempted to correct for such a structural issue by relying on point processes (Schutte, 2016; Zammit-Mangion et al., 2012). Second, examining only conflict onset as a dependent variable disregards variations in violence intensity across the units of analysis, forcing researchers to rely on inferior statistical methods to produce their estimates. Third, the inclusion of spatial lags in the analysis can bias the results if spatial autocorrelation errors between neighbouring grid cells are not accounted for. Fourth, and extracted directly from the preliminary results, the exceptionally large number of grids with zero values are always likely to produce strong significant results. These issues will be addressed in the following sections.

4.5.2 Conflict Intensity

Instead of just relying on the binary dependent variable of conflict onset, the analysis provides a more coherent picture if the intensity of violence is also explored. The *intensity* variable does not just convey if a given grid is in conflict at a specific period, but also how many conflict events occurred in it. The analysis permits a clearer and a more robust explanation of the drivers of violence *vis-á-vis* the binary onset variable. In other words, it explains whether more intense grids are likely to continue escalating or whether violence diffuses into neighbouring grids with time.

Both random and fixed effect spatial models are used to estimate the determinants of conflict intensity. Table 4.2 shows the results of the estimations for both random and fixed effects with the rook model. The first three columns describe the results from the random effect model, while the last three column describe the results from the fixed effects model. columns (1) and (4) show the results without incorporating the spatial lag variable; (2) and (5) with the spatial rook lag and the spatial correlation parameter of the regression errors; (3) and (6) without correcting for spatially correlated errors. It is important to note, that when estimating the full model, the initial value for the spatial correlation parameter is taken to be the estimated ρ from a panel regression with spatially correlated errors based on Baltagi et al. (2007). Analogously, the initial value of λ is the estimated spatial autocorrelation coefficient from the spatial autoregressive model.

Moreover, in addition to the list of explanatory variables used in estimating conflict onset, I include the temporal lagged variable of the number of events associated for each armed group. The armed grouped are aggregated to three categories: opposition forces, regime forces, and unknown forces.

First, in the random effects model, conflict intensity in $t-1$ and $t-2$ are positive and significant for all three specifications. This is only true for violence intensity in $t-1$ in the fixed effect model. Second, if opposition forces were part of the violence in $t-1$, then violence is likely to escalate in the period after. However, if regime forces were reported to be part of the violence in $t-1$ then violence is more likely to de-escalate. These results hold for both the random and fixed effect model, and for all three model specifications. These results shed light on the actor-specific characteristics of the conflict in Syria between the regime and opposition forces. First, violence perpetrated by the regime varies across space from one period to the other, while violence involving the opposition is concentrated in certain spatial pockets. Second, whenever the opposition forces are active in a given cell, then the violence is likely to intensify in the next period.

More importantly, intense violence actions are less likely to spread with the same intensity across neighbouring cells. λ, which denotes the first degree rook spatial coefficient, carries a negative sign. This implies that violent conflict is likely to de-escalate if violence in neighbouring grids intensifies. This result is intuitive, reiterating that conflict

Table 4.2: Spatial determinants of violence intensity - rook model

	(1)	(2)	(3)	(4)	(5)	(6)
Intensity_{t-1}	0.472^{***}	0.397^{***}	0.397^{***}	0.371^{***}	0.369^{***}	0.371^{***}
	(0.023)	(0.026)	(0.026)	(0.028)	(0.026)	(0.026)
Intensity_{t-2}	0.177^{***}	0.051^{***}	0.052^{***}	0.022	0.021	0.022
	(0.015)	(0.015)	(0.015)	(0.016)	(0.015)	(0.015)
Opposition_{t-1}	2.966^{***}	2.592^{***}	2.624^{***}	2.491^{***}	2.426^{***}	2.484^{***}
	(0.247)	(0.246)	(0.247)	(0.255)	(0.244)	(0.245)
Regime_{t-1}	-0.126^{***}	-0.263^{***}	-0.262^{***}	-0.267^{***}	-0.266^{***}	-0.267^{***}
	(0.039)	(0.048)	(0.048)	(0.051)	(0.049)	(0.049)
Unknown_{t-1}	1.978^{***}	1.548^{***}	1.534^{***}	1.417^{***}	1.426^{***}	1.416^{***}
	(0.152)	(0.160)	(0.160)	(0.166)	(0.159)	(0.159)
Population	-0.311^{***}	-0.448^{*}	-0.505^{**}			
	(0.070)	(0.230)	(0.220)			
Wealth	0.015^{***}	0.022^{**}	0.024^{***}			
	(0.003)	(0.010)	(0.009)			
Roads	0.003^{***}	0.007^{***}	0.007^{***}			
	(0.000)	(0.001)	(0.001)			
Country Border	-0.019	-0.022	-0.028			
	(0.015)	(0.049)	(0.047)			
Altitude (SD)	-0.015	-0.026	-0.039			
	(0.011)	(0.038)	(0.036)			
Farmland	0.001^{***}	0.002^{*}	0.002^{*}			
	(0.000)	(0.001)	(0.001)			
λ		-0.041	-0.001^{***}		-0.102^{*}	0.021
		(0.026)	(0.000)		(0.054)	(0.020)
ρ		0.066^{**}			0.128^{**}	
		(0.034)			(0.057)	
R^2	0.923			0.164		
Adj. R^2	0.923			0.093		
Num. obs.	3887	3887	3887	3887	3887	3887
Effects	Random	Random	Random	Fixed	Fixed	Fixed
Spatial Lag	No	Yes	Yes	No	Yes	Yes
Spatial Error	No	Yes	No	No	Yes	No

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors in parentheses. Spatial lags are denoted by λ and are based on the rook model. The spatial error denoted by ρ is based on balatagi. Hausmann test's null hypthesis is rejected for all specifications, therefore the random effect model is inconsistent. R^2 and Adj. R^2 are not used for the spatial model, as the estimation are based on ML. Population in 100,000; Wealth in 1,000: Altitude in 100m, and Number of Towns in 100.

rarely takes place with the same severity across regions, rather is concentrated in specific pockets, and its intensity diffuses to neighbouring regions with time. However, the first degree spatial coefficient is only significant at the 10% level with the inclusion of the spatial error parameter.

In the random effect model, the time invariant explanatory variables show that farmlands and accessible areas have a positive effect on the intensity of violence. In contrast to the results from the conflict onset estimations, wealthy grids increase the propensity of violence escalation, while more populous areas decrease it. This implies that although conflict takes place in more populated grids, it tends to be short-lived with time. The R^2 is equal to 0.923 in the random effect model, which is a significantly higher than 0.164 of the fixed effect model. This raises concerns on the consistency of the estimates coming out of the random effect model, especially that the inclusion of time-invariant variables positively bias the fitness of the model. The null hypothesis of the Hausman specification test is rejected for both the temporal and spatial models, implying that the results from the random effects model are largely invalid.

Table A.5 in the appendix of chapter 4 shows the results using the queen spatial matrix. All other variables are kept similar to the original estimations. The one notable difference is related to the spatial lag. With the exclusion of the error parameter, the spatial coefficient λ is positive and significant at the 10% level. However, the sign changes back to minus as soon as I account for the spatial autocorrelation error parameter ρ. This reconfirms the importance of correcting for plausible spatially autocorrelated errors in analysing conflict event data with spatial dependence. Moreover, the stronger effect of the queen spatial lag coefficient reflects the importance of including adequate spatial boundaries around the grid area to capture better the diffusion of violence.

Predicting Intensity

Next, only predictions of the consistent fixed effect model are presented. The results are shown for both of the in-sample and out-of-sample predictions. Given that the ROC is not appropriate for continuous response variables, the Root-Mean-Squared-Errors (RMSE) and the Mean Absolute Error (MEA) are alternatively used to compare the validity of the predictions. The RMSE and MEA determine the magnitude of the errors of the residuals between the observed and predicted values. They are robust measures of the performance accuracy of the predictions, but are only valid for comparing errors across models but not within each model. The interpretation of the results are intuitive: the closer the RSME and MEA are to zero, the stronger the predictive performance of the model.

Table 4.3: RMSE and MEA for in- and out- of sample predictions of violent conflict escalation

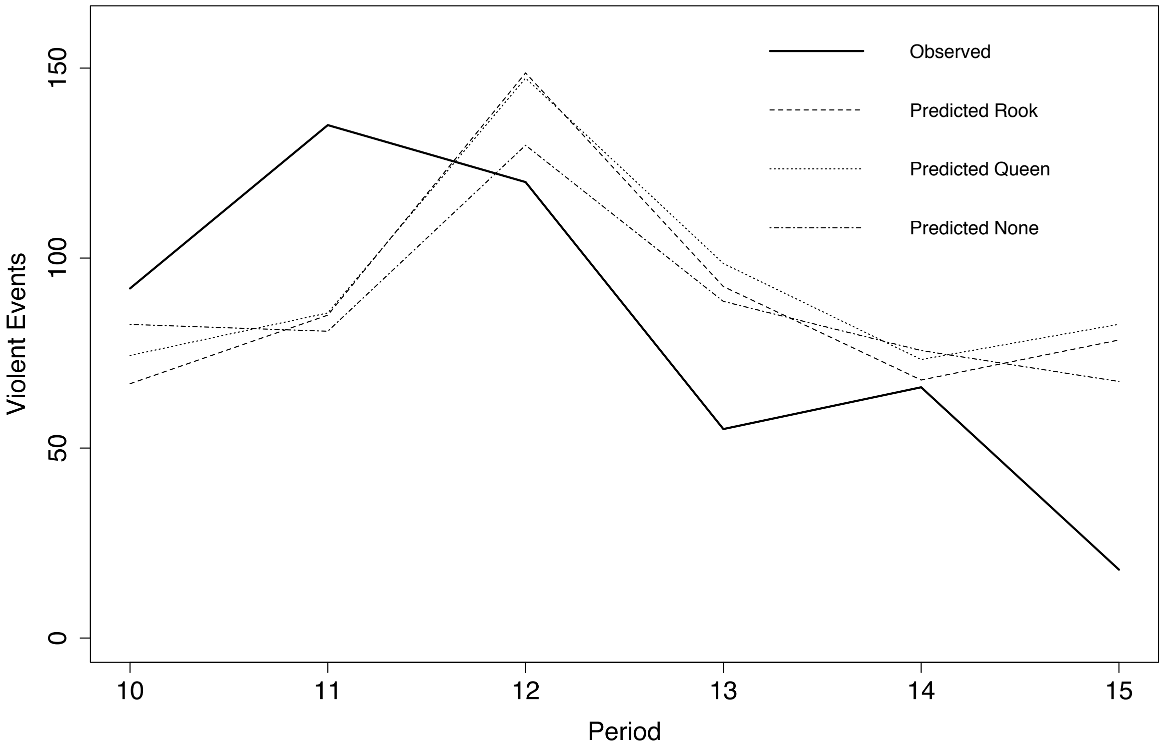
	In-Sample			Out-of-Sample		
RMSE	1.274	1.278	1.289	1.463	1.789	1.786
MEA	0.210	0.219	0.236	0.293	0.317	0.324
Effects	Fixed	Fixed	Fixed	Fixed	Fixed	Fixed
Spatial Lag	No	Rook	Queen	No	Rook	Queen
Spatial Error	-	Yes	Yes	-	Yes	Yes

RMSE denotes Root-Mean-Squared-Error, and MEA denotes Mean-Absolute-Error.

Table 4.3 shows the results for all three specifications: the temporal model, the spatio-temporal using the rook, and the queen cases, respectively. The baseline temporal model without any spatial lags performs better than both spatio-temporal models. This difference in performance is spelled out more clearly in the strict out-of-sample predictions.

However, the difference in performance between the rook and queen models is very marginal, and alternating at times across the in-sample and out-of-sample predictions. This is especially true when using the RMSE measures as an indicator of performance. For example, the predictions using the queen model perform slightly better in the out-of-sample forecasts. In contrast, the model with the rook spatial lag outperforms it when using MEA measures. Given that the MEA structure resembles more accurately the loss function of the model (i.e, event counts falling further away from the mean do not necessarily outweigh ones closer to it), it is safe to conclude that the model with the rook spatial lag performs best in both in-sample and out-of-sample predictions.

Figure 4.4: Out-of-sample forecasts for intensity of violence



Note: *The test data includes periods 10-15. Out-of-sample predictions are extracted from the estimates of the fixed effects spatial model.*

These marginal differences in performance of the temporal fixed effect model over its spatio-temporal counterparts are better depicted in Figure 4.4. The figure shows the event forecasts for the fixed effects models from the out-of-sample predictions. They are plotted against the observed violence intensity for each period in the “test” dataset, which includes the periods 10 to 15. As can be clearly observed from the figure, the model without any spatial lags is able to forecast the actual events best. Moreover, the results show a slightly enhanced performance in forecasting of the rook model over the queen model.

To summarise the results, the inclusion of the first order spatial lag does not enhance the predictive performance of the model in both predictions of violent conflict intensity and onset. This especially becomes more evident when applying the more stringent and robust out-of-sample predictions. The results contrast with that of Weidmann and Ward (2010) and Ward and Gleditsch (2002), who argue that space is an unquestionable predictor of violence. Although I find that violence in neighbouring spatial grids does indeed have a significant impact on violence onset and intensity, it does not necessarily enhance our predictions on future levels of violence. These differences could be largely driven by the initial choice of the unit of analysis, which has a substantial impact on the direction of the results. Further discussion and tests in that regard are presented in the following section

4.6 Robustness Checks and Statistical Corrections

Two methodological considerations are implemented and discussed to ensure the robustness of the results for both predicting conflict onset and escalation. The first issue is related to the choice of the grid size as the unit of analysis, and its impact on the predictive capabilities of the spatio-temporal models. The second issue is related to possibility of existing self-selection based on the grid characteristics between the units of analysis, which might lead to biased estimates.

Grid and Sample Size:
The use of a pre-specified grid cell sizes can significantly impact the outcomes of the estimations. In order to account for any estimations errors resulting from the baseline choice of the grid size, which has an area of 625 km^2 (side equals to 25 km), the estimations are presented for smaller grid sizes to ensure the robustness and validity of the results.

Table 4.4 shows the results of the spatial logit model estimations of the determinants of conflict onset using grid cells of an area of 100 km^2. Naturally, the smaller the size of the grid cells, the higher the number of observations. The panel dataset now contains 24,388 observations from thirteen periods. In all three specification, there are no impactful changes of the results when using smaller grid cells. Even more, the results are strongly

Table 4.4: Determinants of conflict onset - 10 km grid cells

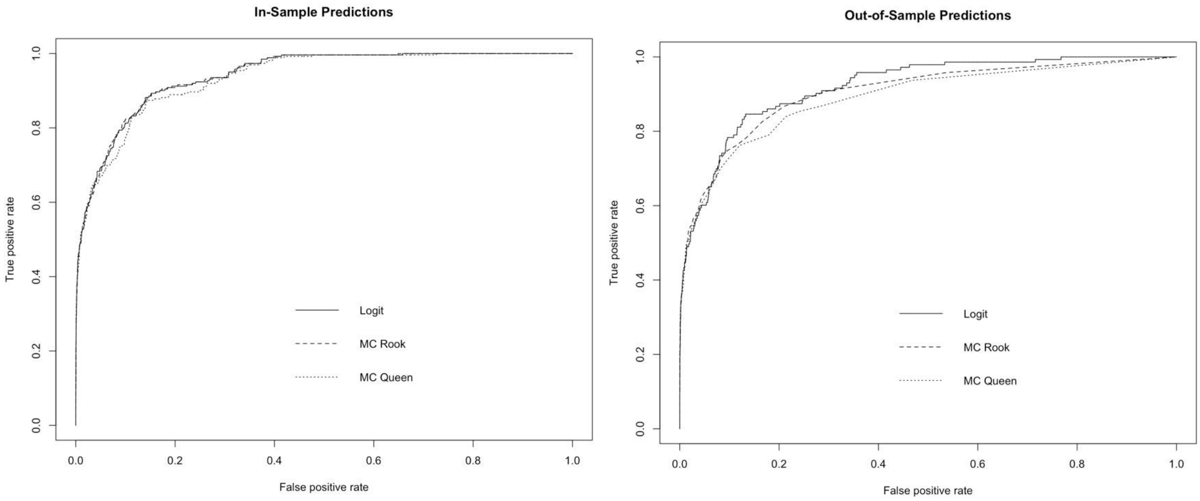
	(1)	(2)	(3)
Conflict_{t-1}	1.99^{***}	1.89^{***}	1.86^{***}
	(0.21)	(0.21)	(0.21)
Conflict_{t-2}	1.68^{***}	1.57^{***}	1.57^{***}
	(0.22)	(0.22)	(0.22)
Population	17.60^{***}	18.22^{***}	18.41^{***}
	(1.97)	(1.96)	(1.94)
Wealth	-0.73^{***}	-0.76^{***}	-0.77^{***}
	(0.08)	(0.08)	(0.08)
Roads	1.66^{***}	1.56^{***}	1.55^{***}
	(0.24)	(0.24)	(0.24)
Altitude (SD)	0.39^{***}	0.35^{***}	0.33^{**}
	(0.13)	(0.13)	(0.13)
Country Border	-0.01	0.03	-0.02
	(0.18)	(0.18)	(0.18)
Farmland	0.98^{***}	0.79^{***}	0.74^{***}
	(0.18)	(0.18)	(0.18)
λ		3.62^{***}	4.94^{***}
		(0.62)	(0.89)
Num. obs.	24388	24388	24388
Model	Logit	MC	MC
Spatial Lag	No	Rook	Queen

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard deviation in parantheses. The first spatial lag is denoted by λ, where specification (2) uses the rook model and (3) uses the queen model. Population in 100,000; Wealth in 1,000: Altitude in 100m; Roads, farms and country border are binary variables.

more significant when using the Monte Carlo estimations with the spatial lag for both the queen and rook models in comparison the baseline models.

Moreover, as shown in figure 4.5 there are also no notable changes in the predictive capabilities of the various models compared to the baseline grid size. Even more, the

Figure 4.5: ROC plots of conflict onset - 10 km grids



***Note**: The Area Under the Curve (AUC) values for the out-of-sample simulated predictions for the temporal, rook spatio-temporal, and queen spatio-temporal specification are 0.920, 0.904, 0.890, respectively.*

queen model performs worse in both in-sample and out-of-sample predictions.[22] The results also extend to larger grid cells of 50km, as well as to estimations of violence escalation.[23]

Sample Selection:

The use of the sample selection models account for the high number of zeros in the dependent variable, which might be inconsistently driving the strong significance of the results. This is true for both conflict onset and conflict intensity. The number of censored grids in all periods is 3,682, which implies that about 94% of the total number of observation are zeros. Therefore any results stemming out of zero driven dependent variable should be dealt with cautiously. However such large number of zero observations are not outside the norm in conflict event data with high spatial precision levels ()

Two types of selection models are widely used from the Tobit-2 Heckman-type sample

---

22 *It is important to note that the spatial lagged models use a smaller number of iterations (500) in generating the MCMC simulations given the sheer size of the dataset.*

23 *Table of results are available upon request.*

standard selection models: the two-step estimation and the maximum likelihood estimation (Heckman, 1979; Amemiya, 1984). The self-selection corrections clarifies which part of the violence intensity is driven by the temporal and spatial explanatory variables, and which part is due to the fact that zero grids are fundamentally different from non-zero grids.

In theory, the use of the sample selection correction would produce better estimates of conflict intensity. However, in practice there are a number of statistical challenges in its implementation for the current scenario. First, when estimating the selection equation, the use of the probit model on the panel data structure does not allow for spatial autocorrelation corrections. The alternative use of the Bayesian estimation of the spatial autoregressive probit model produces consistent estimates. The downside, is that the Inverse Mill's Ratio (IMR) is difficult to accurately compute in that case. The outcome equation can only be properly then estimated with OLS using Heckman's two-stage estimations. This also poses drawbacks on the nature of the panel data, and the lagged temporal variables used in the estimations and predictions. Third, to date, there are no consistent and coherent statistical packages that combine both the spatial panel maximum likelihood estimations, with the Tobit-2 Heckman models. These shortcoming are perhaps mainly driven by the need to correct for spatial autocorrelation at both the selection equation (conflict onset) and the outcome equation (conflict intensity).

Based on these shortcomings, the results of the estimations in table A.6 do not correct for the spatial autocorrelation of the residuals at neither the selection nor the outcome equations. Therefore, the significance of the results should be approached carefully The table A.6 shows the results from outcome equation by running both maximum likelihood (columns 1-3) and two-stage estimations (columns 4-6), with and without the inclusion of the spatial lag (rook and queen). In general, there are no notable changes in the results from the baseline model. Intensity of violence in period $t-1$ is still a strong predictor of conflict escalation in period t. However, the effect is not significant for $t-2$. Population size, access to roads, and border cells are all positive and significant with the

ML estimations. This also true for the spatial lags in both the queen and rook models. Based on the BIC values, the model with queen spatial variable performs best.

4.7 Conclusion

In this chapter, temporal and spatial dependencies are modelled to predict both conflict onset and intensity. The analyses are based on own sample data collected on the war in Syria between March 2011 and May 2013. Conflict events in 15 periods, which are selected randomly from within this time frame, are aggregated to 25 km sided grids. Using spatio-temporal autoligistic and spatial panel maximum likelihood regressions, in-sample and out-of-sample predictions are drawn from Monte Carlo simulations for both conflict onset and escalation.

The analyses suggest that space is a significant determinant of both violent conflict onset and escalation. The likelihood of violence onset in a specific spatial unit is positively affected by the existence of conflict its neighbouring regions. However, the intensification of violence in neighbouring regions is likely to diminish the violence in this spatial unit.

Nevertheless, the inclusion of these spatial lags do not considerably enhance the predictive power of the models. Of course, predicting conflict escalation is much more difficult than that of onset, as the nature of violent intensification is driven by a number of unobservable factors which are difficult to account for in the analysis. Challenges in providing better forecasts are driven by data and methodological limitations. Methodologically, I show that there are still substantial challenges in finding appropriate statistical tools that provide unbiased and consistent estimates. Data-wise, the lack of accurate micro-level data on conflict is still one of the main obstacles to build well-performing forecasts. As seen in chapter 3, conflict data from media sources is insufficient in both its quantity (i.e., representation of real violent events) and the adequacy of its geographic spread.

Future research in violent conflict should turn the focus on providing better representative data at the micro-level and enhanced statistical tools. Detailed high precision spatial and temporal data that can be produced using novel methodologies such as crowdseeding are a crucial leap forward. Based on the findings of this chapter, and as argued by (Cederman and Weidmann, 2017), forecasts with fine-grained spatio-temporal variations are feasible and informative. Therefore, given the right tools and data, researchers are able to provide meaningful predictions that can reliably inform practitioners and development agencies working in violent conflict and fragile settings.

Bibliography

Agresti, A. (1994). Simple capture-recapture models permitting unequal catchability and variable sampling effort. *Biometrics 50*(2), 494–500.

Amemiya, T. (1984). Tobit models: A survey. *Journal of Econometrics 24*(1-2), 3–61.

Baillargeon, S. and L.-P. Rivest (2007). Rcapture: loglinear models for capture-recapture in r. *Journal of Statistical Software 19*(5), 1–31.

Balcells, L., L.-A. Daniels, and A. Escribà-Folch (2016). The determinants of low-intensity intergroup violence: The case of Northern Ireland. *Journal of Peace Research 53*(1), 33–48.

Baliki, G. (2014). Measuring conflict incidence in Syria. In *SIPRI Yearbook: Armaments, Disarmament and International Security.* Oxford University Press.

Ball, P., J. Asher, D. Sulmont, and D. Manrique (2003). How many Peruvians have died? *Unpublished Working Paper*.

Baltagi, B. H., S. H. Song, B. C. Jung, and W. Koh (2007). Testing for serial correlation, spatial autocorrelation and random effects using panel data. *Journal of Econometrics 140*(1), 5–51.

Bellows, J. and E. Miguel (2006). War and institutions: New evidence from Sierra Leone. *The American Economic Review 96*(2), 394–399.

Bellows, J. and E. Miguel (2009). War and local collective action in Sierra Leone. *Journal of Public Economics 93*(11), 1144–1157.

Berrebi, C. and E. F. Klor (2006). On terrorism and electoral outcomes theory and evidence from the Israeli-Palestinian conflict. *Journal of Conflict Resolution 50*(6), 899–925.

Berrebi, C. and E. F. Klor (2008). Are voters sensitive to terrorism? Direct evidence from the Israeli electorate. *American Political Science Review 102*(03), 279–301.

Bertram-Hümmer, V. and G. Baliki (2015). The role of visible wealth for deprivation. *Social Indicators Research 124*(3), 765–783.

Besley, T. and T. Persson (2009). Repression or civil war? *The American Economic Review 99*(2), 292–297.

Blair, R. A., C. Blattman, and A. Hartman (2016). Predicting local violence: Evidence from a panel survey in Liberia. *Unpublished Working Paper*.

Blattman, C. (2009). From violence to voting: War and political participation in Uganda. *American Political Science Review 103*(02), 231–247.

Blattman, C. and E. Miguel (2010). Civil war. *Journal of Economic Literature 48*(1), 3–57.

Bocquier, P. and H. Maupeu (2005). Analysing low intensity conflict in Africa using press reports. *European Journal of Population 21*(2-3), 321–345.

Bozzoli, C., T. Brück, and N. Wald (2013). Self-employment and conflict in Colombia. *Journal of Conflict Resolution 57*(1), 117–142.

Braithwaite, A. and S. D. Johnson (2012). Space–time modeling of insurgency and counterinsurgency in Iraq. *Journal of Quantitative Criminology 28*(1), 31–48.

Brakman, S., H. Garretsen, and M. Schramm (2004). The strategic bombing of german cities during World War II and its impact on city growth. *Journal of Economic Geography 4*(2), 201–218.

Brück, T., M. Di Maio, and S. H. Miaari (2014). Learning the hard way: The effect of violent conflict on student academic achievement. *IZA Discussion Paper: 8543*.

Brück, T., P. Justino, P. Verwimp, A. Avdeenko, and A. Tedesco (2016). Measuring violent conflict in micro-level surveys: Current practices and methodological challenges. *The World Bank Research Observer 31*(1), 29–58.

Bueno de Mesquita, E. and E. S. Dickson (2007). The propaganda of the deed: Terrorism, counterterrorism, and mobilization. *American Journal of Political Science 51*(2), 364–381.

Buhaug, H. (2010). Dude, where's my conflict? LSG, relative strength, and the location of civil war. *Conflict Management and Peace Science 27*(2), 107–128.

Buhaug, H. and S. Gates (2002). The geography of civil war. *Journal of Peace Research 39*(4), 417–433.

Buhaug, H., S. Gates, and P. Lujala (2009). Geography, rebel capability, and the duration of civil conflict. *Journal of Conflict Resolution 53*(4), 544–569.

Buhaug, H. and K. S. Gleditsch (2005). The origin of conflict clusters: Contagion or bad neighborhoods. In *Third European Consortium for Political Research General Conference, Budapest*, Volume 8.

Buhaug, H. and K. S. Gleditsch (2008). Contagion or confusion? Why conflicts cluster in space. *International Studies Quarterly 52*(2), 215–233.

Buhaug, H. and J. K. Rød (2006). Local determinants of african civil wars, 1970–2001. *Political Geography 25*(3), 315–335.

Bühlmann, M. and M. Freitag (2006). Individual and contextual determinants of electoral participation. *Swiss Political Science Review 12*(4), 13–47.

Caruso, R. and E. F. Klor (2012). Political economy studies on the Israeli-Palestinian conflict: Introduction. *Peace Economics, Peace Science, and Public Policy 18*(2), 1–10.

Cederman, L.-E. (2008). Articulating the geo-cultural logic of nationalist insurgency. In S. Kalyvas, I. Shapiro, and T. Masoud (Eds.), *Order, Conflict, and Violence*, pp. 242–270. Cambridge University Press, Cambridge.

Cederman, L.-E. and N. B. Weidmann (2017). Predicting armed conflict: Time to adjust our expectations? *Science 355*(6324), 474–476.

Chao, A. (1987). Estimating the population size for capture-recapture data with unequal catchability. *Biometrics 43*(4), 783–791.

Collier, P. and A. Hoeffler (2004). Greed and grievance in civil war. *Oxford Economic Papers 56*(4), 563–595.

Condra, L. N. and J. N. Shapiro (2012). Who takes the blame? The strategic effects of collateral damage. *American Journal of Political Science 56*(1), 167–187.

Darroch, J. N., S. E. Fienberg, G. F. Glonek, and B. W. Junker (1993). A three-sample multiple-recapture approach to census population estimation with heterogeneous catchability. *Journal of the American Statistical Association 88*(423), 1137–1148.

Davenport, C. (2009). *Media Bias, Perspective, and State Repression: The Black Panther Party.* Cambridge University Press.

De Bono, A. and Chatenoux (2014). A global exposure model for GAR 2015. *UNEP-GRID, GAR 2015 background papers for global risk assessment.*

Dixon, P. M. (2006). *Ripley's K Function.* John Wiley & Sons, Ltd.

Drakos, K. and A. Gofas (2006). The devil you know but are afraid to face: Underreporting bias and its distorting effects on the study of terrorism. *Journal of Conflict Resolution 50*(5), 714–735.

Dunson, D. B. and C. Xing (2009). Nonparametric bayes modeling of multivariate categorical data. *Journal of the American Statistical Association 104*(487), 1042–1051.

Earl, J., A. Martin, J. D. McCarthy, and S. A. Soule (2004). The use of newspaper data in the study of collective action. *Annual Review of Sociology 30*, 65–80.

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

Elhorst, J. P. (2003). Specification and estimation of spatial panel data models. *International Regional Science Review 26*(3), 244–268.

Fearon, J. D. and D. D. Laitin (2003). Ethnicity, insurgency, and civil war. *American Political Science Review 97*(01), 75–90.

Ferguson, N. T. (2015). Just the two of us? Civil conflicts, pro-state militants, and the violence premium. *Terrorism and Political Violence*, 1–27.

Fjelde, H. and D. Nilsson (2012). Rebels against rebels explaining violence between rebel groups. *Journal of Conflict Resolution 56*(4), 604–628.

Ganor, B. (2011). *The Counter-Terrorism Puzzle: A Guide for Decision Makers*. Transaction Publishers.

Gerber, A. S., G. A. Huber, D. Doherty, C. M. Dowling, C. Raso, and S. E. Ha (2011). Personality traits and participation in political processes. *The Journal of Politics 73*(03), 692–706.

Getmansky, A. and T. Zeitzoff (2014). Terrorism and voting: The effect of rocket threat on voting in Israeli elections. *American Political Science Review 108*(03), 588–604.

Geyer, C. J. and E. A. Thompson (1992). Constrained Monte Carlo maximum likelihood for dependent data. *Journal of the Royal Statistical Society*, 657–699.

Gleditsch, K. S., N. W. Metternich, and A. Ruggeri (2014). Data and progress in peace and conflict research. *Journal of Peace Research 51*(2), 301–314.

Goldstone, J. A., R. H. Bates, D. L. Epstein, T. R. Gurr, M. B. Lustik, M. G. Marshall, J. Ulfelder, and M. Woodward (2010). A global model for forecasting political instability. *American Journal of Political Science 54*(1), 190–208.

Gould, E. D. and E. F. Klor (2010). Does terrorism work? *The Quarterly Journal of Economics 125*(4), 1459–1510.

Groseclose, T. and J. Milyo (2005). A measure of media bias. *The Quarterly Journal of Economics 120*(4), 1191–1237.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica 47*(1), 153–161.

Hegre, H., G. Østby, and C. Raleigh (2009). Poverty and civil war events a disaggregated study of Liberia. *Journal of Conflict Resolution 53*(4), 598–623.

Hendrix, C. and H.-J. Brinkman (2013). Food insecurity and conflict dynamics: Causal linkages and complex feedbacks. *Stability: International Journal of Security and Development 2*(2), 1–18.

Hendrix, C. S. and I. Salehyan (2015). No news is good news: Mark and recapture for event data when reporting probabilities are less than one. *International Interactions 41*(2), 392–406.

Hijmans, R. J., L. Guarino, C. Bussink, P. Mathur, M. Cruz, and E. Rojas (2004). DIVA-GIS Database.

Hsiang, S. M., K. C. Meng, and M. A. Cane (2011). Civil conflicts are associated with the global climate. *Nature 476*(7361), 438–441.

Iyengar, S. and K. S. Hahn (2009). Red media, blue media: Evidence of ideological selectivity in media use. *Journal of Communication 59*(1), 19–39.

Jaeger, D. A., E. F. Klor, S. H. Miaari, and M. D. Paserman (2012). The struggle for Palestinian hearts and minds: Violence and public opinion in the Second Intifada. *Journal of Public Economics 96*(3), 354–368.

Jaeger, D. A. and M. D. Paserman (2008). The cycle of violence? An empirical analysis of fatalities in the Palestinian-Israeli conflict. *American Economic Review 98*(4), 1591–1604.

Jaeger, D. A., M. D. Paserman, et al. (2009). The shape of things to come? On the dynamics of suicide attacks and targeted killings. *Quarterly Journal of Political Science 4*(4), 315–342.

Johnston, P. B. (2012). Does decapitation work? Assessing the effectiveness of leadership targeting in counterinsurgency campaigns. *International Security 36*(4), 47–79.

Johnston, P. B. and A. K. Sarbahi (2016). The impact of US drone strikes on terrorism in Pakistan. *International Studies Quarterly 60*(2), 203–219.

Jordan, J. (2009). When heads roll: Assessing the effectiveness of leadership decapitation. *Security Studies 18*(4), 719–755.

Justino, P. (2009). Poverty and violent conflict: A micro-level perspective on the causes and duration of warfare. *Journal of Peace Research 46*(3), 315–333.

Justino, P., T. Brück, and P. Verwimp (2013). Micro-level dynamics of conflict, violence, and development: A new analytical framework. In *A Micro-Level Perspective on the Dynamics of Conflict, Violence, and Development*. Oxford University Press.

Justino, P., T. Brück, and P. Verwimp (2016). New developments in measuring the welfare effects of conflict exposure at the micro-level. In D. Becker, R. Bhavnani, and P. Huth (Eds.), *Peace and Conflict 2016,*. Routledge.

Kalyvas, S. (2005). Warfare in civil wars. In I. Duyvesteyn and J. Angstrom (Eds.), *Rethinking the Nature of War*, pp. 88–108. Routledge.

Kalyvas, S. N. (2004). The urban bias in research on civil wars. *Security Studies 13*(3), 160–190.

Kalyvas, S. N. (2008). Ethnic defection in civil war. *Comparative Political Studies 41*(8), 1043–1068.

Kalyvas, S. N., P. Lange, R. H. Bates, E. Comisso, P. Hall, J. Migdal, and H. Milner (2006). *The Logic of Violence in Civil War*. Cambridge University Press.

Kocher, M. A., T. B. Pepinsky, and S. N. Kalyvas (2011). Aerial bombing and counterinsurgency in the Vietnam War. *American Journal of Political Science 55*(2), 201–218.

Krueger, A. B. and J. Malečková (2003). Education, poverty and terrorism: Is there a causal connection? *The Journal of Economic Perspectives 17*(4), 119–144.

Landman, T. and A. Gohdes (2013). A matter of convenience: Challenges of non-random data in analyzing human rights violations during conflicts in Peru and Sierra Leone. In T. Seybolt, J. Aronson, and B. Fischhoff (Eds.), *Counting Civilian Casualties: An Introduction to Recording and Estimating Nonmilitary Deaths in Conflict.*, pp. 77–92. Oxford University Press.

Leetaru, K. and P. A. Schrodt (2013). Gdelt: Global data on events, location, and tone, 1979–2012. In *ISA Annual Convention*, Volume 2. Citeseer.

Linke, A. M., F. D. Witmer, E. C. Holland, and J. O'Loughlin (2017). Mountainous terrain and civil wars: Geospatial analysis of conflict dynamics in the Post-Soviet Caucasus. *Annals of the American Association of Geographers 107*(2), 520–535.

Linke, A. M., F. D. Witmer, and J. O'Loughlin (2012). Space-time granger analysis of the war in Iraq: A study of coalition and insurgent action-reaction. *International Interactions 38*(4), 402–425.

Lyall, J. (2009). Does indiscriminate violence incite insurgent attacks? Evidence from Chechnya. *Journal of Conflict Resolution 53*(3), 331–362.

Manrique-Vallier, D. (2016). Bayesian population size estimation using dirichlet process mixtures. *Biometrics 72*(4), 1246–1254.

Manrique-Vallier, D., M. E. Price, and A. Gohdes (2013). Multiple systems estimation techniques for estimating casualties in armed conflicts. In T. Seybolt, J. Aronson, and B. Fischhoff (Eds.), *Counting Civilian Casualties: An Introduction to Recording and Estimating Nonmilitary Deaths in Conflict.* Oxford University Press.

Manrique-Vallier, D. and J. P. Reiter (2014). Bayesian estimation of discrete multivariate latent structure models with structural zeros. *Journal of Computational and Graphical Statistics 23*(4), 1061–1079.

Metternich, N. W., C. Dorff, M. Gallop, S. Weschle, and M. D. Ward (2013). Anti-government networks in civil conflicts: How network structures affect conflictual behavior. *American Journal of Political Science 57(4)*, 892:911.

Miaari, S. H. and R. M. Sauer (2011). The labor market costs of conflict: Closures, foreign workers, and Palestinian employment and earnings. *Review of Economics of the Household 9*(1), 129–148.

Miguel, E., S. Satyanath, and E. Sergenti (2004). Economic shocks and civil conflict: An instrumental variables approach. *Journal of Political Economy 112*(4), 725–753.

Millo, G. and G. Piras (2012). splm: Spatial panel data models in R. *Journal of Statistical Software 47*, 1–38.

Minoiu, C. and O. N. Shemyakina (2014). Armed conflict, household victimization, and child health in Côte d'Ivoire. *Journal of Development Economics 108*, 237–255.

Mullainathan, S. and A. Shleifer (2005). The market for news. *American Economic Review 95*(4), 1031–1053.

O'Loughlin, J. and F. D. Witmer (2011). The localized geographies of violence in the North Caucasus of Russia, 1999–2007. *Annals of the Association of American Geographers 101*(1), 178–201.

Pape, R. A. (2014). *Bombing to Win: Air Power and Coercion in War*. Cornell University Press.

Price, B. C. (2012). Targeting top terrorists: How leadership decapitation contributes to counterterrorism. *International Security 36*(4), 9–46.

Price, M., A. Gohdes, and P. Ball (2015). Documents of war: Understanding the Syrian conflict. *Significance 12*(2), 14–19.

Raleigh, C. and H. Hegre (2009). Population size, concentration, and civil war: A geographically disaggregated analysis. *Political Geography 28*(4), 224–238.

Raleigh, C., A. Linke, H. Hegre, and J. Karlsen (2010). Introducing acled: An armed conflict location and event dataset special data feature. *Journal of Peace Research 47*(5), 651–660.

Reed, W. (2000). A unified statistical model of conflict onset and escalation. *American Journal of Political Science 44*(1), 84–93.

Rivest, L.-P. and S. Baillargeon (2007). Applications and extensions of Chao's moment estimator for the size of a closed population. *Biometrics 63*(4), 999–1006.

Rosendorff, B. P. and T. Sandler (2004). Too much of a good thing? The proactive response dilemma. *Journal of Conflict Resolution 48*(5), 657–671.

Rustad, S. C. A., H. Buhaug, Å. Falch, and S. Gates (2011). All conflict is local: Modeling sub-national variation in civil conflict risk. *Conflict Management and Peace Science 28*(1), 15–40.

Salehyan, I. and K. S. Gleditsch (2006). Refugees and the spread of civil war. *International Organization 60*(02), 335–366.

Schelling, T. C. (1980). *The Strategy of Conflict*. Harvard University Press.

Schutte, S. (2016). Regions at risk: Predicting conflict zones in African insurgencies. *Political Science Research and Methods*, 1–19.

Stoll, D. (1993). *Between Two Armies in the Ixil Towns of Guatemala*. Columbia University Press.

Sundberg, R. and E. Melander (2013). Introducing the UCDP georeferenced event dataset. *Journal of Peace Research 50*(4), 523–532.

Tinsley, H. E. and D. J. Weiss (1975). Interrater reliability and agreement of subjective judgments. *Journal of Counseling Psychology 22*(4), 358.

Tollefsen, A. F., H. Strand, and H. Buhaug (2012). PRIO-GRID: A unified spatial data structure. *Journal of Peace Research 49*(2), 363–374.

Tranchant, J.-P., P. Justino, C. Müller, et al. (2014). Political violence, drought and child malnutrition: Empirical evidence from andhra pradesh, india. *Households in Conflict Network Working Paper No 173*.

Trelles, A. and M. Carreras (2012). Bullets and votes: Violence and electoral participation in Mexico. *Journal of Politics in Latin America 4*(2), 89–123.

Van der Windt, P. and M. Humphreys (2016). Crowdseeding in Eastern Congo: Using cell phones to collect conflict events data in real time. *Journal of Conflict Resolution 60*(4), 748–781.

Verwimp, P., P. Justino, and T. Brück (2009). The analysis of conflict: A micro-level perspective. *Journal of Peace Research 46*(3), 307–314.

Voors, M. J., E. E. Nillesen, P. Verwimp, E. H. Bulte, R. Lensink, and D. P. Van Soest (2012). Violent conflict and behavior: A field experiment in Burundi. *The American Economic Review 102*(2), 941–964.

Ward, M., N. Metternich, C. Dorff, M. Gallop, F. Hollenbach, A. Schultz, and S. Weschle (2013). Learning from the past and stepping into the future: Toward a new generation of conflict prediction. *International Studies Review 15*(4), 473–490.

Ward, M. D. and K. S. Gleditsch (2002). Location, location, location: An MCMC approach to modeling the spatial context of war and peace. *Political Analysis 10*(3), 244–260.

Ward, M. D., B. D. Greenhill, and K. M. Bakke (2010). The perils of policy by p-value: Predicting civil conflicts. *Journal of Peace Research 47*(4), 363–375.

Weidmann, N. B. (2011). Violence “from above” or “from below”? The role of ethnicity in bosnia’s civil war. *The Journal of Politics 73*(04), 1178–1190.

Weidmann, N. B. (2016). A closer look at reporting bias in conflict event data. *American Journal of Political Science 60*(1), 206–218.

Weidmann, N. B. and M. D. Ward (2010). Predicting conflict in space and time. *Journal of Conflict Resolution 54*(6), 883–901.

Wood, E. J. (2003). *Insurgent Collective Action and Civil War in El Salvador*. Cambridge University Press.

Wood, R. M. (2010). Rebel capability and strategic violence against civilians. *Journal of Peace Research 47*(5), 601–614.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*. MIT press.

Woolley, J. T. (2000). Using media-based data in studies of politics. *American Journal of Political Science 44*(1), 156–173.

Yared, P. and G. Padró i Miquel (2012). The political economy of indirect control. *Quarterly Journal of Economics 127*(2), 947–1015.

Zammit-Mangion, A., M. Dewar, V. Kadirkamanathan, and G. Sanguinetti (2012). Point process modelling of the Afghan war diary. *Proceedings of the National Academy of Sciences 109*(31), 12414–12419.

Zanotti, J. (2010). *Israel and Hamas: Conflict in Gaza (2008-2009)*. DIANE Publishing.

Zhukov, Y. M. (2012). Roads and the diffusion of insurgent violence: The logistics of conflict in Russia’s North Caucasus. *Political Geography 31*(3), 144–156.

Zhukov, Y. M. and M. A. Baum (2016). Reporting bias and information warfare. *Unpublished Working Paper*.

Appendices

Appendix of Chapter 2

Table A.1: Differences in individual and household characteristics between indirect exposed and control groups

	Neighbourhood Damage Major			Neighbourhood Damage All		
	No	Yes	p-value	No	Yes	p-value
N	1479	518		625	1372	
Age	35.02	35.07	0.951	34.82	35.13	0.653
Gender (Female)	48.1	50.2	0.443	47.0	49.4	0.357
Governorate % (freq)			<0.001			<0.001
North Gaza	15.5	28.4		9.3	23.2	
Gaza	33.9	38.0		28.3	38.0	
Deir El-Balah	15.4	15.8		18.9	14.0	
Khan Yunis	22.7	9.7		31.2	13.9	
Rafah	12.5	8.1		12.3	10.9	
Voted E 2006	0.42	0.47	0.117	0.48	0.41	0.026
Participated 2006	0.67	0.66	0.632	0.67	0.66	0.745
Income			0.223			0.405
No income	66.8	68.1		65.3	68.0	
< 1200 NIS	15.7	12.5		15.3	14.7	
1200-2000 NIS	9.9	12.1		10.1	10.6	
2000-3000 NIS	5.8	6.3		7.4	5.2	
> 3000 NIS	1.9	1.1		1.9	1.5	
Education			0.739			0.471
No Level	11.8	12.8		10.7	12.6	
Elementary	12.6	10.6		12.5	11.9	
Basic	23.9	23.2		23.4	23.9	
Secondary	32.2	34.0		32.0	33.0	
Tertiary	19.4	19.3		21.4	18.5	
HH Size	6.59	6.52	0.640	6.67	6.53	0.347
Employed	73.9	74.9	0.685	73.8	74.3	0.839
N child	2.82	2.86	0.704	2.80	2.84	0.722

The exact fisher test is used instead of the Chi squared test to account for the low number of observations within some category of the following variables: Governorate, income, and education.

Table A.2: Impact of direct and indirect violence on voting moderate - multinomial approach

	Own Dwelling Damage				Neighborhood Dwelling Damage			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Vote M 2006	5.39***	4.95***	5.32***	5.36***	5.32***	5.19***	5.31***	5.47***
	(0.32)	(0.41)	(0.31)	(0.32)	(0.31)	(0.52)	(0.31)	(0.37)
Damage All	0.75***	0.24			0.38**	0.15		
	(0.17)	(0.43)			(0.18)	(0.45)		
Vote M 2006*Damage All		1.07				0.18		
		(0.68)				(0.65)		
Damage Major			1.56**	2.54**			−0.00	0.30
			(0.65)	(1.02)			(0.19)	(0.46)
Vote M 2006*Damage Major				8.92***				−0.58
				(0.84)				(0.68)
Gender (Female)	−0.42**	−0.44**	−0.42**	−0.42**	−0.47**	−0.47**	−0.45**	−0.45**
	(0.20)	(0.20)	(0.20)	(0.20)	(0.20)	(0.20)	(0.20)	(0.20)
Education (Basic)	−0.61*	−0.65*	−0.56	−0.53	−0.61*	−0.61*	−0.60*	−0.60*
	(0.35)	(0.35)	(0.35)	(0.35)	(0.35)	(0.35)	(0.35)	(0.35)
Education (Secondary)	−0.58*	−0.62*	−0.57*	−0.55	−0.61*	−0.61*	−0.59*	−0.59*
	(0.35)	(0.35)	(0.35)	(0.35)	(0.35)	(0.35)	(0.35)	(0.35)
Income (2000-3000 NIS)	0.82*	0.82*	0.80*	0.80*	0.93**	0.91*	0.88*	0.87*
	(0.46)	(0.46)	(0.46)	(0.46)	(0.47)	(0.47)	(0.47)	(0.47)
Income Hamas	−1.74***	−1.77***	−1.79***	−1.82***	−1.75***	−1.76***	−1.75***	−1.75***
	(0.36)	(0.36)	(0.36)	(0.37)	(0.36)	(0.36)	(0.36)	(0.36)
Emotional Instability	0.02	0.02	0.03**	0.03**	0.02**	0.02**	0.03**	0.03**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
BIC	4319.68	4397.51	4333.55	4414.33	4315.82	4396.37	4321.75	4399.98
Log Likelihood	−1862.28	−1856.56	−1869.22	−1864.97	−1860.54	−1856.21	−1863.51	−1858.01
Num. obs.	1701	1701	1701	1701	1693	1693	1693	1693

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. Standard Errors in Parenthesis. voting Extreme is used as the baseline category in the multinomial regression analysis for both the depedenet variable and the Election variable of 2006. All other sub-categories (Not Participate, Others, Undecided) are not included in the table. Moreover, only significant income and education categories are shown. Employment, Number of children, and the constant term are repressed.

Table A.3: Impact of direct and indirect damage on election participation

	Own Dwelling Damage				Neighborhood Dwelling Damage			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Partic 2006	0.64^{***}	0.69^{***}	0.63^{***}	0.63^{***}	0.63^{***}	0.71^{***}	0.63^{***}	0.63^{***}
	(0.07)	(0.10)	(0.07)	(0.07)	(0.07)	(0.12)	(0.07)	(0.08)
Damage All	−0.05	0.01			−0.04	0.02		
	(0.07)	(0.11)			(0.07)	(0.11)		
Partic 2006*Damage All		−0.10				−0.11		
		(0.13)				(0.15)		
Damage Major			0.01	−0.22			−0.04	−0.06
			(0.23)	(0.32)			(0.07)	(0.12)
Partic 2006*Damage Major				0.47				0.02
				(0.48)				(0.15)
Age	-0.01^{***}	-0.01^{***}	-0.01^{***}	-0.01^{***}	-0.01^{***}	-0.01^{***}	-0.01^{***}	-0.01^{***}
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Gender (Female)	−0.14	−0.14	-0.18^{*}	−0.14	−0.14	−0.14	−0.14	−0.14
	(0.08)	(0.08)	(0.07)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)
HH Size	0.04^{**}	0.04^{**}		0.04^{**}	0.04^{*}	0.04^{*}	0.04^{*}	0.04^{*}
	(0.02)	(0.02)		(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
N child	-0.05^{*}	-0.05^{*}		-0.05^{*}	-0.05^{*}	-0.05^{*}	-0.05^{*}	-0.05^{*}
	(0.02)	(0.02)		(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
AIC	1993.49	1994.98	2017.28	1995.01	1976.54	1978.00	1976.62	1978.59
BIC	2091.74	2098.69	2093.86	2098.71	2074.70	2081.62	2074.78	2082.21
Log Likelihood	-978.75	-978.49	-994.64	-978.50	-970.27	-970.00	-970.31	-970.30
Deviance	1957.49	1956.98	1989.28	1957.01	1940.54	1940.00	1940.62	1940.59
Num. obs.	1734	1734	1755	1734	1726	1726	1726	1726

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table A.4: Impact of direct and indirect damage on voting extreme - subsample

	Own Dwelling Damage				Neighborhood Dwelling Damage			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Vote E 2006	2.38***	2.25***	2.39***	2.38***	2.37***	2.30***	2.37***	2.39***
	(0.12)	(0.16)	(0.12)	(0.12)	(0.12)	(0.20)	(0.12)	(0.14)
Damage All	−0.16	−0.32			0.02	−0.04		
	(0.12)	(0.19)			(0.12)	(0.19)		
Vote E 2006*Damage All		0.28				0.11		
		(0.24)				(0.25)		
Damage Major			−1.04*	−3.99			−0.00	0.04
			(0.50)	(111.53)			(0.13)	(0.21)
Vote E 2006*Damage Major				3.01				−0.07
				(111.53)				(0.27)
Income Hamas	0.68***	0.69***	0.70***	0.70***	0.69***	0.69***	0.69***	0.69***
	(0.20)	(0.20)	(0.21)	(0.21)	(0.20)	(0.20)	(0.20)	(0.20)
Emotional Instability	−0.02*	−0.02*	−0.02*	−0.02*	−0.02*	−0.02*	−0.02*	−0.02*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
HH Size	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
N child	−0.04	−0.04	−0.04	−0.04	−0.04	−0.04	−0.04	−0.04
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
AIC	606.44	607.11	603.63	605.46	606.31	608.13	606.34	608.27
BIC	692.72	698.18	689.91	696.53	692.57	699.18	692.60	699.33
Log Likelihood	-285.22	-284.55	-283.81	-283.73	-285.16	-285.06	-285.17	-285.14
Deviance	570.44	569.11	567.63	567.46	570.31	570.13	570.34	570.27
Num. obs.	892	892	892	892	891	891	891	891

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Appendix of Chapter 3

Codebook

The codebook includes all definitions and detailed information of all events recorded in DISC conflict-event dataset on Syria. We divide the coded events into three main categories: (A) Violent events, (B) other forcible events, and (C) political and organisational events.

Definition of an Event: *"An occurrence that is related to violence and the political conflict in Syria; either an instance of the direct use of violence or other forcible actions, protests, or related internal and external political events."*

An instance is only coded as a single event if:

- It takes place at the same day;
- It takes place at the same place;
- It is of the same event type;
- It involves at least two of the same actors throughout.

Unless all the above conditions hold, an instance is coded as multiple events. The following events are included in the Syrian Event Database:

A. Violent Events

Definition of a Violent Event: *"An event that is related to harmful force through the use of any manufactured arms or the use of unmanufactured arms (e.g. sticks, stones, fire, etc) in a manner that results in bodily harm to another person or destruction or damage to buildings or infrastructure."*

The violent event is subdivided according to: (I.) the form of violence, (II.) the tools, and (III.) the outcome in terms of deaths, injuries and/or destruction.

I. Forms of violent events

There are three forms of violent events

i. **Violence between armed actors:** use of violence by an organised armed actor against another organised armed actor.[24]

ii. **Military offensive** in a town or area: use of violence projected onto a physical infrastructure in some section of a territory.

iii. **Violence against civilians:** use of violence by an organised armed actor against civilians.

II. Tools of violent events

The tools include a list of manufactured weapons and is divided into the 9 categories. The tools are meant to capture the extent of arms that is used within each event. It is possible to select more than one means for the same event.

i. **Aerial attack:** any explosives dropped or ammunition fired from an airborne aeroplane or helicopter.

ii. **Anti-craft attack:** any explosives or ammunition fired, aimed at an airborne target.

iii. **Shelling and heavy arms:** firing of ground-to-ground explosive devices. (tanks, mortar, etc)

iv. **Small and medium arms:** portable fire arms (RPGs, automatic and semi-automatic guns, snipers)

---

[24] *Organised armed actor: a group of people having announced a name and some purpose for their group, and having on at least one occasion used armed force to achieve any of the groups stated purposes. Includes Syrian Government military forces.*

v. **Stationary explosive devices:** explosives that are set off either from a distance, or upon contact with a victim (e.g. mine, improvised explosive devices, car bombs).

vi. **Suicide bombing:** perpetrator setting off explosives with the intent of harming other individuals, in such close proximity to the perpetrator that he/she does not hope to survive it.

vii. **Chemical weapons or BCW:** chemical or biological weapons (e.g. white phosphorus)

viii. **Other arms:** Any other hand-held manufactured or unmanufactured arms not being fire arms.

ix. **Other, unknown**

III. Outcome of violent events

The outcomes of a violent event are divided into four categories. Destruction and Damage of buildings, death of civilians or armed actors, injury of civilians or armed actors, and territory gains or losses. It is possible to select more than one outcome for the same event.

i. **Destruction** of or damage to a building, a piece of cultural property or to infrastructure.

- **Dwellings**: a place of residence
- **Cultural property**: a building or object which is considered to be, for secular or religious reasons, of importance to archaeology, prehistory, history, literature, art or science.
- **Government building**: a building in which central or local non-military government officials are based (e.g. ministry, town hall)
- **A secular public building**: not being cultural property or a government

building (e.g. a school, hospital)

- **A religious public building**: not being cultural property (e.g. a mosque, church)
- **Transportation** infrastructure (e.g. a road, bridge, airport)
- **Infrastructure** relating to water, power utilities, or communication (e.g. a power plant, dam, a radio mast)
- **An Army building or facility** e.g. army base, military school or airport

ii. Death

- Death of one or more civilians
- Death of one or more members of an organised armed actor
- Death of person, unknown.

iii. Bodily Harm

- Bodily harm to one or more civilians
- Bodily harm to one or more members of an organised armed actor
- Bodily harm to a person, unknown.

iv. Territory gain by an organised armed actor.[25]

B. Other Forcible Events

Definition of Other Forcible events: *"events that do not involve the use of violence action as defined above. These actions are sub-categorised as follows into 6 forms: (i)*

[25] *Territory gain is defined as the takeover of an area by an organized armed actor. The armed actors retreating from the gained area included to keep record of the military territorial changes.*

forced displacement, (ii) kidnapping and detainment, (iii) organised theft and looting, (iv) sexual violence (v) road and border closures, and (vi) utility cuts."

I. Direct and Indirect Forced Displacement

Direct Forced Displacement: *"the forcing away from their place of habitual residence, of civilians by an organised armed actor, by means of expulsion or other coercive acts directly targeting said civilians."*

Expulsion: *"removal, by (threat of) force, from the physical structure one lives in."*

Indirect Forced Displacement: *"civilians leaving their places of habitual residence to in order to avoid being subject to armed force, without said civilians being expulsed or directly coerced through violence by an organized armed actor or because it is not possible to fulfil basic needs."*

Additional information on displacement also includes the form of the forced displacement should be noted where applicable; that is external or internal; where "Internal Displacement" signifies that citizens who are forced away from their place of habitual residence remain within Syria, and "External Displacement" signifies that citizens who are forced away from their place of habitual residence leave Syria. Moreover, the country of destination in case of a forced external displacement, and the destination city, town, or district in case of forced internal displacement is also coded. Lastly, the total number of displaced civilians per event is entered where available.

Providing this added eases the association of event type (i.e., violence) and means that actually force people to flee there country rather than settle within. We also would like to observe that citizens fleeing their country are located closer to the boarders than those internally displaced.

II. Kidnapping and Detainment

Kidnapping: *"carrying away and detaining an unarmed civilian by force."*

Detainment: *"confining of one or more member(s) of an armed actor or protestor(s), by another organised armed actor, irrespective of whether the confined individuals are formally recognised as Prisoners of War."*

If kidnapped civilians, detainees, or prisoners of war are subjected to violence, two separate events are coded.

III. Theft and Looting

Theft and Looting: *"taking without permission, of: (a) Cultural property: as defined above; (b) Private property: the property of civilians, excluding cultural property. (c) Public property: property owned by the state or community, excluding cultural property."*

IV. Sexual Violence

Sexual Violence: *"forcing, through any means, of an individual, male or female, into performing sexual acts without the consent of said individual."*

V. Road or Border Closures

Road closures: *"prohibiting of the use of a road by a temporary installation and/or deployment of (a) member(s) of an organised armed actor of roads, within Syrian borders by an organized armed actor."*

Border closures: *"blocking or closure of border checkpoint between Syria and its neighbouring countries, by an organized armed actor."*

VI. Utility Cuts

Utility Cuts: *"intended cuts of water, electricity, fuel/gas, or communication supply (internet, telephone, etc) by an organised actor through means other than the use of violence."*

C. Protests, Political and Organisational Events

I. Protests and Riots

Protests: *“a group of organised or unorganised civilians expressing discontent with the government or an organised armed actor in a public space, in a manner not involving the use of violence.”*

Riots: *“a group of organised or unorganised civilians expressing discontent with the government or an organised armed actor in a public space, in a manner involving the use of violence.”*

If violence was used against protesters by an organised armed actor, two separate events are coded. Moreover, if riots are accompanied with theft for example, two separate events are also entered.

II. Internal Political Events

Internal Political events: *“are actions by internal political actors involved directly or indirectly in the conflict that can affect its course. Internal political actors are political representatives of the Syrian government or the opposition forces, or political representatives of any armed actor that are involved directly or indirectly in conflict.”*

Internal political events are divided into three independent, yet interlinked categories:

i. **Agreements:** “all decisions and promises reached within or between political internal actor regarding the Syrian conflict, even if their implementation does not take place.”

ii. **Meetings:** “all meetings held by political internal actors that involve matters related to the Syrian conflict; may it be to support the military or reaching a consensus to cease-fire. ”

iii. **Statements:** “all official statement issued by political internal actors that either accuse, condemn, or announce any new developments in relation to the Syrian conflict.”

Example	Events Coded
Representative of group x met with representative of group y to find a solution to bring in aid to Homs, and they issued the following statement:	Two events: 1. Meeting 2. Statement
Assad met with his cabinet to discuss heightening security measures in Damascus.	One Event: Meeting
The opposition forces agreed to form a temporary government after their meeting in Istanbul. x, y, z attended the talks.	two events: Agreement; Meeting actors: x, y, z. If not stated then simply opposition

III. External Political Events

External Political events: *"are actions by external political actors involved directly or indirectly in the conflict that can affect its course."*[26]

These events are divided in a similar fashion to internal political events.

i. **Agreements:** all promises made and decisions reached between at least two political actors (where at least one of them is an external political actor) regarding the Syrian conflict, even if their implementation does not take place.

ii. **Meetings:** all meetings held by political external actors that involve matters related to the Syrian conflict; may it be to support the military or reaching a consensus to cease-fire.

iii. **Statements:** all official statement issued by external political actors that either accuse, condemn, or announce any new developments in relation to the Syrian conflict.

IV. Defections and Resignations

Defections: *"official announcements by a single member or a group of members of an*

[26] *External political actors are representatives of international community or countries or international representative of countries that are involved directly or indirectly in conflict.*

organized armed actor of leaving the armed actor for another."

Resignations: *"official announcements by a single member or a group of members of an organised armed actor of leaving the organised armed actor without joining another."*

V. Formations, Disbandment, and Alliances

Formations: *"official announcements by an unorganized group of armed actors in the formation of an organized group of armed actor."*[27]

Disbandment: *"official announcements of splintering of an organized group of armed actors into several armed actors or to an unorganized group of armed actors."*

Alliances: *"official announcements of the formation of an organized group of armed actors from different groups of organized armed actors."*

Event Actors

The actors included in the Syrian Event Database are divided into the following categories

Armed Actors: 1. Syrian Arab Forces 2. Shabiha 3. Free Syrian Army (FSA) 4. Syrian Liberation Front (SLF) 5. Syrian Islamic Front (SIF) 6. Jabhat Al-Nusra 7. Kurdish Democratic Union Party 8. Hezbollah 9. Other, specified 10. Other, unspecified, regime 11. Other, unspecified, opposition 12. Not Reported 13. Unknown

Unarmed Actors: 1. Civilians 2. Protestors Unorganised civilians or protestors are considered one actor.

Internal Political Actors: 1. Syrian Government 2. Bashar Al-Assad (President) 3. Syrian National Council 4. National Coalition for Syrian Revolutionary and Opposition

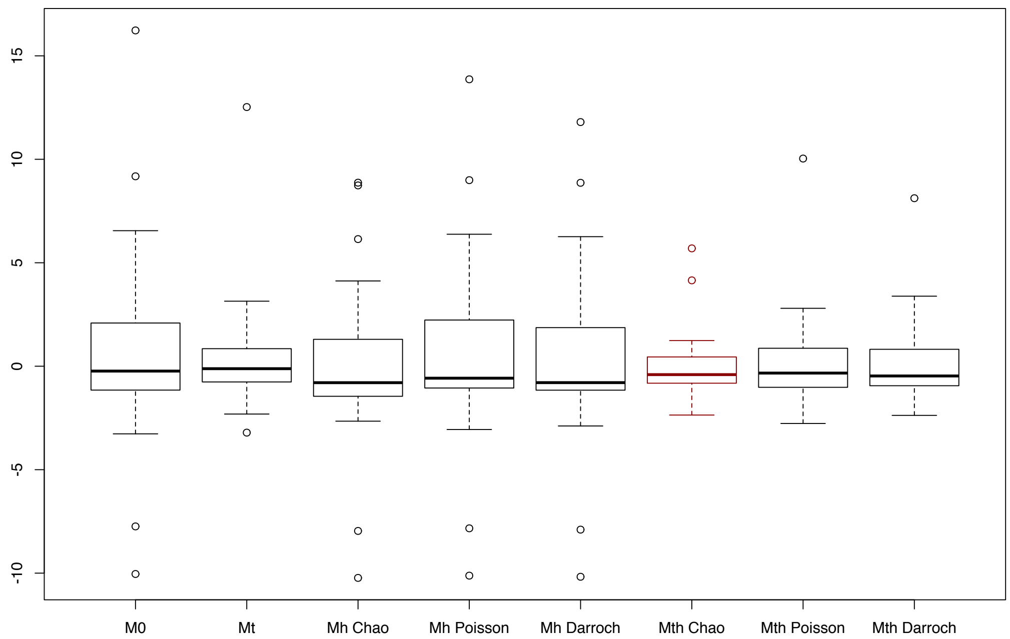
[27] *Unorganized armed actors are a group of people that do not have a name for their group and do not function under the command of another armed actor, where they have at least once engaged in an armed action.*

Forces 5. Local Coordination Committees of Syria 6. Kurdish Supreme Committee 7. Muslim Brotherhood 8. Popular Front for Change and Liberation 9. Other, specified 10. Other, unspecified

External Political Actors: 1. United States 2. Iran 3. Russia 4. Turkey 5. Jordan 6. Lebanon 7. Israel 8. European Union 9. Cooperation Council for the Arab States of the Gulf 10. Arab League 11. UN Security Council/General Assembly 12. Syrian representative at the UN (Bashar Al-Jaafari) 13. Kofi Annan 14. Akhdar El-Ibrahimi 15. Other, Specified

End of Codebook

Figure A.1: Boxplot of pearson residuls of capture-recapture models



Appendix of Chapter 4

Table A.5: Spatial determinants of violence intensity - queen model

	(1)	(2)	(3)	(4)	(5)	(6)
Intensity$_{t-1}$	0.472***	0.395***	0.397***	0.371***	0.365***	0.370***
	(0.023)	(0.026)	(0.026)	(0.028)	(0.026)	(0.026)
Intensity$_{t-2}$	0.177***	0.050***	0.052***	0.022	0.021	0.022
	(0.015)	(0.015)	(0.015)	(0.016)	(0.015)	(0.015)
Opposition$_{t-1}$	2.966***	2.680***	2.623***	2.491***	2.535***	2.482***
	(0.247)	(0.246)	(0.247)	(0.255)	(0.241)	(0.245)
Regime$_{t-1}$	−0.126***	−0.262***	−0.261***	−0.267***	−0.263***	−0.266***
	(0.039)	(0.048)	(0.048)	(0.051)	(0.048)	(0.049)
Unknown$_{t-1}$	1.978***	1.590***	1.533***	1.417***	1.462***	1.418***
	(0.152)	(0.160)	(0.160)	(0.166)	(0.158)	(0.159)
Population	−0.311***	−0.324	−0.520**			
	(0.070)	(0.246)	(0.220)			
Wealth	0.015***	0.017	0.025***			
	(0.003)	(0.010)	(0.009)			
Roads	0.003***	0.007***	0.007***			
	(0.000)	(0.001)	(0.001)			
Country Border	−0.019	−0.011	−0.028			
	(0.015)	(0.051)	(0.047)			
Altitude (SD)	−0.015	−0.003	−0.042			
	(0.011)	(0.039)	(0.036)			
Farmland	0.001***	0.002*	0.002*			
	(0.000)	(0.001)	(0.001)			
λ		−0.099***	0.009		−0.252***	0.058**
		(0.037)			(0.081)	(0.027)
ρ		0.174***			0.305***	
		(0.043)			(0.067)	
R^2	0.923			0.164		
Adj. R^2	0.923			0.093		
Num. obs.	3887	3887	3887	3887	3887	3887
Effects	Random	Random	Random	Fixed	Fixed	Fixed
Spatial Lag	No	Yes	Yes	No	Yes	Yes
Spatial Error	No	Yes	No	No	Yes	No

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors in parentheses. Spatial lags are denoted by λ and are based on the queen model. The spatial error denoted by ρ is based on balatagi. Hausmann test's null hypthesis is rejected for all specifications, therefore the random effect model is inconsistent. R^2 and Adj. R^2 are not used for the spatial model, as the estimation are based on ML. Population in 100,000; Wealth in 1,000: Altitude in 100m, and Number of Towns in 100.

Table A.6: ML and two-stage Heckman selection models - determinants of violence escalation

	(1)	(2)	(3)	(4)	(5)	(6)
Intensity_{t-1}	0.38^{***}	0.39^{***}	0.38^{***}	0.42^{***}	0.42^{***}	0.42^{***}
	(0.11)	(0.11)	(0.11)	(0.13)	(0.13)	(0.13)
Intensity_{t-2}	0.02	0.02	0.02	−0.02	−0.02	−0.02
	(0.05)	(0.05)	(0.05)	(0.06)	(0.06)	(0.06)
Opposition_{t-1}	1.01	1.68^{**}	1.48^{*}	2.95^{***}	2.99^{***}	2.96^{***}
	(0.70)	(0.82)	(0.78)	(1.12)	(1.12)	(1.12)
Regime_{t-1}	−0.22	−0.21	−0.20	−0.31	−0.31	−0.31
	(0.19)	(0.21)	(0.20)	(0.25)	(0.25)	(0.25)
Unknown_{t-1}	1.00^{*}	1.20^{**}	1.18^{**}	1.91^{**}	1.93^{**}	1.91^{**}
	(0.53)	(0.61)	(0.58)	(0.80)	(0.80)	(0.80)
Population	0.37^{***}	0.33^{***}	0.33^{***}	0.12	0.12	0.12
	(0.08)	(0.08)	(0.08)	(0.09)	(0.09)	(0.09)
Roads	0.01^{*}	0.00	0.01	0.01	0.01	0.01
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Country Border	1.23^{**}	1.14^{*}	1.19^{*}	0.09	0.15	0.13
	(0.60)	(0.64)	(0.64)	(0.86)	(0.86)	(0.87)
Altitude (SD)	0.47	0.29	0.34	0.12	0.11	0.14
	(0.43)	(0.46)	(0.45)	(0.63)	(0.63)	(0.63)
Farmlands	0.01^{***}	0.01^{**}	0.01^{**}	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)
Intensity_{s1}		0.16^{*}	0.30^{**}		−0.05	0.02
		(0.09)	(0.13)		(0.16)	(0.23)
Model	ML	ML	ML	2Stage	2Stage	2Stage
Spatial Lag	No	Rook	Queen	No	Rook	Queen
Num. obs.	3887	3887	3887	3887	3887	3887
Censored	3682	3682	3682	3682	3682	3682
Observed	205	205	205	205	205	205
BIC	2280	2274	2267			
R^2				0.53	0.53	0.53

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$, Standard errors in parentheses.