Critical configurations

Settlement patterns and ethnic violence

Diss. ETH No. 18257

Critical Configurations: Settlement Patterns and Ethnic Violence

A Dissertation Submitted to ETH ZURICH
in Partial Fulfillment of the Requirements for the Degree

Doctor of Sciences

Presented by Nils Benedikt Weidmann
Master of Comparative and International Studies, ETH Zurich
Diplom-Informatiker, Albert-Ludwigs-Universität Freiburg i. Br.

Born October 11, 1976
Citizen of Germany

Accepted upon the recommendation of
Prof. Lars-Erik Cederman, ETH Zurich, Switzerland
Prof. Håvard Hegre, University of Oslo, Norway
Prof. Michael D. Ward, University of Washington, Seattle, USA

2009

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

TO MY DEAR WIFE REGINE AND OUR BABY DAUGHTER, LISKA.

ACKNOWLEDGMENTS

During the research for this dissertation, I received help, support and encouragement from many people. As my dissertation advisor, Lars-Erik Cederman supported this project from its very beginning, and his input was crucial in bringing my research to the present stage. No less importantly, however, Lars-Erik encouraged me to make the move from my original field of study (computer science) to political science, a step that has turned out to be extremely exciting.

I benefitted a lot from the interaction with my co-advisors, Håvard Hegre and Michael D. Ward. Håvard Hegre made it possible for me to visit the International Peace Research Institute in Oslo (PRIO) a number of times, and provided thorough feedback on my project. My research stay with Michael D. Ward at the University of Washington had a tremendous influence on this dissertation, not only because I was able to rely on his expertise in spatial analysis. Moreover, I am greatly indebted to him for all the support and advice he offered besides this project.

The International Conflict Research group at the Center for Comparative and International Studies was a stimulating place to carry out my research. My current and former colleagues Christa Deiwiks, Lutz Krebs, Patrick Kuhn, Thomas Sattler, Livia Schubiger, Sebastian Schutte, Judith Vorrath, Stefanie Walter and Julian Wucherpfennig provided numerous helpful comments on my work. Luc Girardin shared some of his impressive programming expertise, which helped me solve quite a few computational problems in no time.

Other people contributed to this project. Thanks to Jan Ketil Rød, I was introduced to the secrets of GIS. The skills I learned from him had a great influence on this dissertation project. I presented earlier versions of the dissertation chapters at various conferences, and received invaluable feedback and comments from Halvard Buhaug, Erik Gartzke, Kristian Gleditsch, Nils Petter Gleditsch, Simon Hug, Dominic Johnson, Vally Koubi, Erik Melander, David Meyer and Brian Pollins.

Not least, I would like to thank Elin Arbin for thorough linguistic editing of the manuscript.

I gratefully acknowledge the funding of this project through a grant by ETH Zurich (Research Grant TH -4/05-3).

Zurich, February 2009
Nils B. Weidmann

ABSTRACT

Quantitative research on ethnic conflict has tended to neglect causal mechanisms. By primarily resorting to macro-level indicators and tests, existing studies fail to give a sufficiently precise account of the actors and their interactions that ultimately lead to conflict. This dissertation shows that a detailed analysis of ethnic settlement patterns can help us get closer to the group processes during conflict. With the help of Geographic Information Systems (GIS), computational modeling and spatial statistics, the analysis presented here provides new insights into the role of group geography as a determinant for conflict.

The dissertation is organized around three core questions. The first focuses on ethnic groups as the level of analysis and examines how geographic concentration influences the risk of a group being involved in conflict. Two competing explanations exist: a motivation-driven mechanism which claims that concentrated groups are more likely to fight for their territory, and an opportunity-driven mechanism in which concentrated groups are more prone to collective mobilization because of facilitated social interaction. However, no systematic attempt has been made to test the explanatory impact of these mechanisms. Using detailed GIS data on group settlement regions, my analysis shows that the relationship of concentration and conflict hinges on the increased interaction opportunities in concentrated groups.

With the help of a spatial computational model, this dissertation aims to provide a more detailed account of group mobilization in a spatial context. The model simulates the spread of mobilization through group populations, relying on the assumption that direct social interaction is important for an individual's decision to participate. Consequently, the mobilization influence from one location to the other is proportional to the number of people traveling between these locations. The model computes an indicator for the proportion of a group that can be mobilized given the parameters of the model. Groups with high mobilization values should thus be more susceptible for conflict, a hypothesis which is confirmed in a regression analysis.

The second core question of the dissertation increases the spatial resolution and examines how group geography affects the location of violence. By employing high-resolution geographic data on ethnicity and spatially linked reports of violent events from Bosnia, this dissertation sheds new light on the dynamics of civil wars. I show that the application of violence is driven by territorial ambitions of groups. During conflict, groups aim to create and consolidate larger areas of ethnic homogeneity. Violence is likely at a location if more than one group considers this location to be part of its territory. The analysis provides a detailed perspective on different types of ethnic violence by distinguishing between two-sided violence between the groups' armed factions, and one-sided violence against civilians, both of which seem to be strongly influenced by territorial ambitions.

The literature on group settlement patterns and conflict has treated group geography as an exogenous determinant of conflict. However, research on international conflicts tells us that

territory is as much a cause of conflict as it is a consequence. The third core question applies this logic to an intrastate setting and analyzes changes in group geography that occur as a result of conflict. Comparing settlement patterns in Bosnia before and after the war, the results show that changes in the territorial configuration occur in a strategic way, such that they reduce the territorial contestation between groups. However, my analysis can only partly explain how the occurrence of violence on the ground relates to these systemic changes. Whereas the occurrence of military confrontations seems to inhibit large changes towards ethnic homogeneity, there is no direct relationship between the targeting of civilians and ethnic unmixing.

ZUSAMMENFASSUNG

Die quantitative Forschung zu ethnischen Konflikten ist von einer unzureichenden Spezifikation kausaler Mechanismen gekennzeichnet. Weil sich diese Forschung in erster Linie auf Indikatoren und Tests auf der Makro-Ebene stützt, kann sie keine ausreichend genaue Erklärung geben, wie die beteiligten Akteure und ihre Interaktionen schlussendlich zu Konflikt führen. Diese Dissertation zeigt, dass eine detaillierte Untersuchung der Siedlungsmuster ethnischer Gruppen zum Verständnis der Gruppenprozesse beitragen kann, die sich in bewaffneten Konflikten abspielen. Unter Zuhilfenahme von geographischen Informationsystemen (GIS), Computermodellierung und räumlicher Statistik erzielt die hier vorgestellte Analyse neue Erkenntnisse über den Zusammenhang zwischen der Geographie ethnischer Gruppen und Konflikt.

Die Dissertation ist entlang von drei Kernfragen ausgerichtet. Die erste Kernfrage konzentriert sich auf ethnische Gruppen als Analyseeinheit und untersucht, wie die räumliche Konzentration einer Gruppe das Konfliktrisiko beeinflusst. Es gibt zwei mögliche Erklärungen: Die eine besagt, dass konzentrierte Gruppen eher für ihr Territorium kämpfen und Konzentration somit eine Motivation für Konflikt darstellt. Die andere Erklärung nimmt an, dass konzentrierte Gruppen aufgrund besserer Interaktionsmöglichkeiten der Gruppenmitglieder leichter für Konflikt mobilisiert werden können. Bisher wurde jedoch kein systematischer Versuch gemacht, diese Mechanismen gegeneinander zu testen. Meine Analyse zeigt unter Verwendung von GIS-Daten ethnischer Siedlungsgebiete, dass der zweite Mechanismus empirisch bestätigt wird und somit der Effekt von räumlicher Konzentration einer verbesserten sozialen Interaktion innerhalb einer Gruppe zugeschrieben werden kann.

Mit der Hilfe eines räumlich referenzierten Computermodells führe ich eine genauere Untersuchung der Mobilisierung einer Gruppe im räumlichen Kontext durch. Das Modell simuliert die Ausbreitung von Mobilisierung in Gruppen unter der Annahme, dass direkte soziale Interaktion die Entscheidung eines Individuums über die Teilnahme am Konflikt beeinflusst. Der soziale Einfluss zwischen zwei Orten wird als proportional zur Anzahl der Personen angenommen, welche sich zwischen diesen Orten bewegen. Auf dieser Grundlage berechnet das Modell den Anteil der Gruppe, der sich auf diese Weise mobilisieren lässt. Gruppen mit höheren Werten dieses Indikators sollten eine höhere Konfliktanfälligkeit aufweisen. Diese Hypothese wird in einer Regressionsanalyse bestätigt.

Die zweite Kernfrage der Dissertation erhöht die räumliche Auflösung und untersucht den Einfluss der Verteilung ethnischer Gruppen darauf, wo ethnische Gewalt auftritt. Durch die Verwendung detaillierter GIS-Daten über Ethnizität und räumlich referenzierter Meldungen über Konfliktereignisse aus dem Krieg in Bosnien erzielt meine Analyse neue Erkenntnisse über die Dynamik in Bürgerkriegen. Ich zeige, dass der Ort ethnischer Gewalt durch territoriale Bestrebungen ethnischer Gruppen bestimmt wird. Während eines Konfliktes versuchen die Gruppen, zusammenhängende, ethnisch homogene Regionen zu schaffen. Dies führt zu Gewalt an einem

bestimmten Ort, falls mehr als eine Gruppe diesen Ort als Teil ihres Territoriums ansieht. Ich präsentiere eine detaillierte Analyse von verschiedenen Arten ethnischer Gewalt: Auseinandersetzungen zwischen bewaffneten Gruppen, und Gewalt gegen die Zivilbevölkerung. Wie meine Ergebnisse zeigen, sind beide Arten von Gewalt durch territoriale Bestrebungen bestimmt.

Die bisherige Forschung über Geographie und Konflikt betrachtete die räumliche Anordnung ethnischer Gruppen als exogen. Die Literatur zu internationalen Konflikten jedoch zeigt, dass ein Territorium sowohl ein Auslöser als auch ein Ergebnis von bewaffneten Auseinandersetzungen sein kann. Die dritte Kernfrage dieser Dissertation untersucht dies im Kontext innerstaatlicher Konflikte und analysiert Änderungen im Siedlungmuster einer Gruppe, welche sich als Resultat von ethnischer Gewalt ergeben. Der Vergleich von Siedlungsmustern vor und nach dem Bosnien-Krieg zeigt, dass Änderungen in der Geographie ethnischer Gruppen in einer Weise auftreten, dass sie territoriale Auseinandersetzungen von Gruppen reduzieren. Allerdings kann meine Studie nur teilweise erklären, welche Rolle die lokale Anwendung von Gewalt in diesem systemischen Prozess spielt. Während das Auftreten militärischer Konfrontationen zu geringeren Veränderungen der ethnischen Verteilung führt, finde ich keinen direkten Zusammenhang zwischen Gewalt gegen die Zivilbevölkerung und ethnischer Entmischung.

Contents

Chapter 1

Introduction

1.1 Background

In the past century, humankind was struck by previously unseen dimensions of violent conflict. Within a few decades, the two world wars caused an exorbitant number of casualties, reaching a geographic scope of destruction larger than any conflict before. Since then, however, the number of interstate conflicts has declined steadily. According to recent numbers, there was not a single interstate conflict active in 2007.[1] Has the world become a more peaceful place? Unfortunately, it has not. The trend we have observed over the last decades is that states fight less *against each other*. At the same time, however, there has been a tremendous increase in violence *within states* during the post-World War II period. Media reports from civil wars in Bosnia, Rwanda and Chechnya remind us that what has changed is not the level of political violence in the world, but merely its scope. In 2007, the number of active civil wars reached 29, most of which are unlikely to be resolved in the near future. The world is not a peaceful place – yet.

Even though limited in scope as compared to international wars, civil wars have disastrous consequences, both during and after the conflict. For example, the war in Angola (1975–2002) caused roughly 1.5 million war deaths (Lacina and Gleditsch 2005). The social impact of this figure becomes even more troubling due to the fact that only about 11% of these were battle-related. Clearly, civil wars affect the civilian population in a serious way and with long-lasting consequences. For example, one consequence is the displacement of people. Refugee flows caused by internal conflict are a major challenge for war-torn countries, and the resulting humanitarian concerns are often well beyond what the international community can adequately

[1] Source: *UCDP/PRIO Armed Conflict Dataset* (Harbom, Melander and Wallensteen 2008; Gleditsch et al. 2002)

respond to (United Nations High Commissioner for Refugees 2006). Furthermore, the economic legacy of civil war makes it difficult for countries to recover. Once the investors have left a country, the state budget has been spent on military equipment, and corruption is omnipresent, a country can be caught in what World Bank economist Paul Collier (2003) has called a "conflict trap" – a vicious cycle of violence and economic downturn.

In order to mitigate or prevent conflict, we require an understanding of its causes and dynamics. Much of the current literature on civil wars tends to conclude that civil war is a problem of the poor. Indeed, the correlation between poverty and conflict is high, making one think that conflict is yet another problem developing countries face besides hunger, unemployment, diseases and low education. However, recent events right on Europe's doorstep suggest that this is only partly true. In the early 1990s, the former Yugoslav Federation – certainly not a developing country – experienced a series of violent confrontations between its ethnic groups. Serbs, Croats, Bosniaks, Macedonians and Albanians were caught in recurring cycles of conflict, which eventually led to the complete disintegration of the Federation. Killings reached levels on par with events during World War II, something Europe thought it would never experience again.

Conflicts in Yugoslavia, but also in other wealthy countries like Spain or Northern Ireland, challenge the link between economic poverty and civil war. Instead, they draw our attention to another important dimension of conflict: identity. In fact, ethnic identity features prominently as an explanation for why civil wars occur. Recent findings provide evidence for this: The majority of intrastate wars are fought along ethnic lines (Fearon and Laitin 2003), and ethnic conflicts tend to be more difficult to resolve (Doyle and Sambanis 2000; Licklider 1995). However, despite the seemingly strong relationship between ethnicity and conflict, recent statistical analysis at the country level find ethnic factors to be only a weak explanation of civil war onset. As a result, the quantitative literature assigns only a minor importance to ethnic factors and largely dismisses ethnicity as an explanation for civil war.

The question about whether ethnicity matters continues to be an ongoing debate in the scholarly community. Underlying this question is a methodological issue: Can cross-national studies that operationalize ethnicity by means of a single indicator of ethnic diversity be taken as a sufficiently detailed test of the mechanisms leading groups to violence? Many scholars suggest that they cannot. As Sambanis puts it,

> the already significant gap between the micro-level theories and their macro-level implications is magnified when the micro-macro relationships are studied solely through cross-national statistical analyses. Such studies often overlook information about the causal pathways that link individual or group behavior with the outbreak

of civil war (Sambanis 2004*a*, p. 259).

The second part of Sambanis' quote summarizes the gap in the literature that this dissertation aims to fill. If ethnicity features so prominently in internal conflicts, but we still do not know *how*, would it not be natural to study the relevant actors in these conflicts – the ethnic groups – rather than relying on aggregate indicators at the state level?

1.2 Subject of Research

In the current quantitative literature, the study of ethnic conflict suffers from a dramatic lack of specification of causal actors and mechanisms. Research has concentrated on demographic indicators at the state level rather than giving sufficient attention to the conflict actors and the ethnic groups these actors claim to represent. How do groups and group members coordinate for conflict? Who is the target of violence? In order to further our understanding of how social actors organize and apply violence, recent work on civil war has increased the resolution of the unit of analysis. Rather than comparing states with respect to their risk of civil war, *disaggregated* research explicitly assumes a sub-national perspective by focusing e.g. on ethnic groups, their political position, or the location of conflict. In doing so, we can get closer to the processes of conflict by opening the "black box" between indicators at the state level and the observed resulting conflicts.

The disaggregation of ethnic conflict can be done in various ways. For example, one promising avenue for research is a *conceptual* disaggregation by moving from aggregate ethnic indicators to the level of ethnic groups and conflict actors. Alternatively, *temporal* disaggregation analyzes the sequence of conflict events over time, rather than measuring civil war at the country-year level as many studies do. This dissertation makes a strong case for a third type of disaggregation: the analysis of groups and conflict from a *geographic* perspective. More precisely, I study how conflict is related to settlement patterns of ethnic groups, i.e. the spatial distribution of a group in a country. Why study settlement patterns? First, a group's settlement pattern offers theoretical insights. Group territory is the social space that structures the interactions of its individuals. This social space is where a group's collective mobilization for conflict takes place, so the group's geography partly determines whether or not group mobilization will be successful and result in conflict. Besides their impact on conflict risk in general, settlement patterns can also have an impact on where violence is applied, once inter-ethnic hostilities have started: What particular locations are worth fighting for?

Second, these theoretical insights raise important questions for conflict mitigation and prevention. If the territorial arrangement of groups has an impact on the likelihood and location

of violence, could an artificial *change* in the territorial arrangement limit or even prevent human suffering? In fact, some authors have already proposed territorial solutions to the problem of ethnic violence (see e.g. Kaufmann 1996). Other scholars, however, have argued strongly against these proposed solutions (Sambanis 2000; Laitin 2004). And still, the core problem is that without a detailed understanding of the mechanisms linking group geography and violence, one should be careful in advocating remedies with far-reaching consequences. This work attempts to shed more light on the relationship between conflict and geography in order to help assess the proposed territorial solutions.

My dissertation examines the relationship between settlement patterns and ethnic conflict under three core questions. First, how do settlement patterns affect the risk of a group being involved in violent conflict? Second, how do settlement patterns determine where violence is applied, and third, how does conflict change the ethnic map? The following sections give an introduction to these questions and point the reader to the relevant dissertation chapters.

1.2.1 Settlement Patterns and the Risk of Conflict

Ethnic groups worldwide display considerable variation with regard to their involvement in violent conflict. The field of conflict research tries to explain why some groups experienced significant levels of violence, whereas the majority remain at peace. The first part of my dissertation relates conflict involvement of groups to the characteristics of their settlement pattern. The basic argument is that the geographic arrangement of a group provides interaction opportunities for the group members. If these opportunities are such that the group can efficiently organize for collective action, there is a higher likelihood of conflict. Figure 1.1 illustrates this first core question graphically. The approach here is to measure characteristics of the settlement pattern that facilitate the group's collective mobilization. In a statistical analysis, we can then assess the effect of these characteristics on the risk of conflict involvement.

The first question relates to the role of group geography in conflict more generally. Some authors argue that group territory is important because it constitutes the motivation why groups fight. Quite the contrary, other authors argue that the spatial distribution of group members primarily affects the groups ability to mobilize for collective action. Chapter 4 describes the theoretical background of this motivation vs.opportunity debate and aims to give an answer to it. I show that the distance between the population clusters of a group strongly affects the risk of conflict, whereas a measure for the group's territorial fragmentation has no effect. This result provides support for the opportunity-driven mechanism of group geography.

Although the correlational study I present in Chapter 4 captures the settlement pattern at a high level of detail, it still leaves much to be desired for a test of the mobilization mechanisms

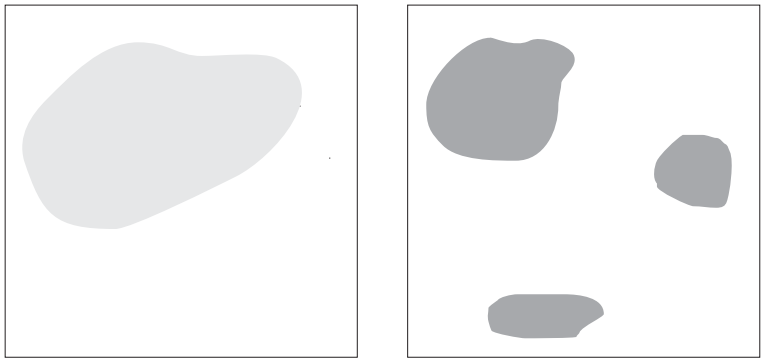


Figure 1.1: First core question: Comparing the settlement pattern of different groups, shown in dark and light grey. Which settlement pattern is more conducive to collective mobilization?

that supposedly drive the relationship between settlement patterns and conflict. In Chapter 5, I attempt to further specify these mechanisms and validate them empirically. For this exercise, I employ a computational model. The model operates directly on spatial data on ethnic groups and simulates how mobilization for collective action spreads through the group population. In doing so, the model computes an indicator for mobilization pervasiveness – the proportion of all group members that can be mobilized for collective action within a certain number of time steps. In a statistical analysis, I compare the simulated mobilization rate to observed conflict involvement and show that groups that are more easily able to mobilize because of their geographic arrangement indeed face a higher risk of conflict.

1.2.2 Settlement Patterns and the Location of Violence

The second core question of this dissertation examines the impact of settlement patterns on the *location* of violence, rather than conflict risk in general as described in the previous part. Here, I move from the focus on settlement patterns of individual groups to intermingled settlement areas of many groups. Figure 1.2 illustrates this question. Again, it shows two groups (dark and light grey), in different settlement configurations relative to each other. I examine whether certain locations along an inter-ethnic territorial boundary are more susceptible for event of violence (shown as stars in Figure 1.2).

I argue that the application of violence during conflict is driven by an ethnic group's aspirations to create larger, coherent territorial clusters of its kin. Locations where territorial claims of different groups clash should see more violence as groups struggle for control of a certain territory. The concurrent claims of different groups to a particular location is what I call territorial *contestation* between groups. In Chapter 6, I study the dynamics of group geography and conflict in Bosnia using pre-war data on ethnic population shares at the municipality level.

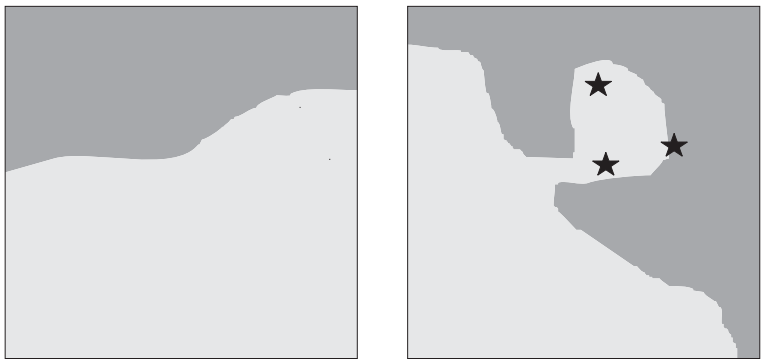


Figure 1.2: Two ethnic groups (dark and light grey), in different boundary configurations (left and right). Under the second core question of this dissertation, it is examined whether and how certain territorial configurations of groups foster the application of violence (shown as stars) at particular locations.

These data are combined with information on conflict events from the *Armed Conflict Location and Events Dataset* (ACLED). My results show that there are indeed higher levels of violence at territorially contested locations.

1.2.3 The Endogeneity of Settlement Patterns and Conflict

Whereas under the first and second perspective outlined above I treat a group's settlement pattern as exogenous to conflict, the third and last core question of this dissertation examines the link back from conflict to changes in the settlement pattern. Chapter 7 builds directly on the previous one, which relates violence to territorial ambitions of ethnic groups. How does the application of violence change the ethnic map? Figure 1.3 again illustrates this question graphically.

Figure 1.3: Third core question of this dissertation: How does the application of violence change the ethnic map? More precisely, if ethnic violence occurs at a particular location (left), does this lead to a consolidation of the ethnic boundaries (right)?

Chapter 7 analyzes the feedback from conflict to changes in the settlement pattern at two levels. First, I take a systemic perspective and compare the pre- and postwar ethnic map. Similar to the previous chapter, this study again examines the war in Bosnia. I find that during the war, there was a significant ethnic segregation of the Bosnian population. However, this unmixing was also strategic in a sense that after the war, clashing claims of groups to particular locations were much less frequent. Second, I try to explain the ethnic population changes at the local level as a result of the application of violence on the ground. Here, the argument is that locations with particularly high levels of violence should see more drastic population changes. However, this hypothesis is not confirmed by empirical evidence.

1.3 Key Contributions

This dissertation makes a number of key contributions to the literature on ethnic violence. In the following, I distinguish between conceptual and methodological issues.

Conceptual: Disaggregating Ethnic Groups The literature on ethnic conflict has had a tendency to treat ethnic groups as unitary actors. This dissertation relates settlement patterns to group mobilization and therefore moves below the group level to intra-group mechanisms of collective action. I draw on literature on social movements, which provides rich accounts of group processes during conflict. However, this literature has not had much intersection with the quantitative work on ethnic conflict. Using computational modeling, I attempt to integrate insights from social movements research to the quantitative study of ethnic conflict.

Conceptual: Disaggregating Ethnic Violence There has been little systematic work on the location of violence in ethnic conflicts. This dissertation presents a framework that relates the application of ethnic violence to the groups' territorial ambitions. In a situation of violent conflict between groups, locations that are of strategic importance to more than one group should be more prone to violence. In conceptualizing violence as a means to satisfy groups' ambitions, I disaggregate violence not only spatially, but also distinguish between different types of violence: two-sided violence between military factions, and one-sided violence against civilians. These two types of violence are likely to follow a different logic, but existing research has lumped them together under a single category.

Conceptual: Endogeneity of Settlement Patterns The relationship between settlement patterns and conflict has only been studied in a single causal direction: How do settlement patterns

affect conflict risk? However, this neglects the reverse causal link. Conflict has a profound impact on the spatial distribution of groups, either by moving a great number of people, or by the systematic extinction of populations. This dissertation presents a first attempt to assess the effect of conflict on settlement patterns by comparing pre- and post-conflict spatial distributions of groups.

Methodological: Geographic Information Systems (GIS) Software For the study of group settlement patterns, this dissertation utilizes recent innovations in the computational processing of spatial data. So-called Geographic Information Systems (GIS) make it possible to collect, represent and process electronic maps of group locations. Although they have been frequently used in geography and related disciplines, GIS are only starting to be introduced to the social sciences and political science in particular. Obviously, the application of GIS to the study of group settlement patterns has many benefits. For the purpose of this dissertation, a global GIS dataset on the locations of ethnic groups has been created, which makes it possible to compute new spatial indicators for a group's settlement pattern. Furthermore, GIS in combination with spatial statistics allows for the spatial disaggregation of ethnic violence as presented in this dissertation.

Methodological: Agent-based Modeling In relying almost exclusively on macro-level correlations between variables, recent quantitative research on ethnic conflict suffers from a significant lack of precision regarding the social mechanisms that bring about these regularities at the macro-level. For a test of the proposed micro-level mechanisms between group geography and conflict mobilization, this dissertation uses agent-based models (ABM). Originally inspired by developments in computer science, ABM is slowly finding its way into the social sciences. ABMs feature a population of social agents living in an artificial space. These agents interact with each other and their environment and allow the researcher to "grow" the observed macro-phenomena from the hypothesized interactions at the micro-level. Unlike many ABMs in the literature, the model presented in this dissertation uses a real-world geography and can thus combine GIS and ABM techniques effectively.

Chapter 2

Groups, Geography and Conflict

THIS chapter introduces the theoretical foundations for the work presented in this dissertation. I start with a definition of ethnic conflict and the "groups" as the main actors involved. I then review the classical literature on ethnic conflict. These works have focused extensively on the motivational roots of conflict: in other words, they can tell us *why* people fight. However, this literature portrays ethnic groups as unitary actors, brushing aside the heterogeneous internal structure of a group. How do some groups manage to mobilize support among many members of their population, and how do they organize violence? This chapter argues that my main *explanans*, group settlement patterns, can provide an answer to this question. I aim to pinpoint the intra-group mechanisms of ethnic conflict by disaggregating ethnic groups and ethnic violence spatially. Studying civil war from a disaggregated perspective is a recent trend in the literature, and I provide an overview of the research and explain the different motivations behind it. I then turn to the relationship between group settlement patterns and conflict. First, I argue that settlement patterns provide interaction opportunities *within* groups in the process of conflict mobilization. Second, settlement patterns also determine interaction opportunities *between* groups, which determine the location of ethnic violence.

2.1 Ethnic Groups in Civil Wars

In the political science literature, the term "civil war" is typically used to denote a conflict that is *internal* to a particular state. More precisely, civil war occurs within the territory of a single state, and only involves actors pertaining to that state. The actor constellation in civil wars is such that a government faces an *internal challenger*, i.e. an opposition group that employs violence to further its goals (Singer and Small 1972; Small and Singer 1982; Fearon and Laitin 2003; Sambanis 2004*b*). Even though the term "internal war" is more precise, it is normally

used interchangeably with "civil war".

Ethnic civil wars constitute a subset of civil wars in which the internal challenger is defined along ethnic lines. A commonly adopted definition of ethnic wars is given by the State Failure Task Force, which defines ethnic wars as "episodes of violent conflict between governments and national, ethnic, religious, or other communal minorities (ethnic challengers) in which the challengers seek major changes in their status" (Marshall, Gurr and Harff 2008). Other scholars have employed similar criteria. For example, Kaufmann (1996) states that "ethnic conflicts are disputes between communities, which see themselves as having distinct heritages, over the power relationship between the communities" (p. 138). Even though these definitions give the reader a general idea about the phenomenon this dissertation examines, it is worth highlighting the differences between ethnic civil war and other forms of ethnic violence.

Ethnic Civil War vs. Violent Protest In its initial stages, violent protest against the government does not constitute ethnic civil war. For example, a spontaneous attack of a village population against the local police station would not qualify as a civil war, unless it is part of a larger, organized effort of a group to achieve a political goal. According to common definitions of civil war, the challenger to the government must be formally organized. For example, the UCDP/PRIO *Armed Conflict Dataset* only codes a series of violent events as civil war if the opposition group has announced a name for itself (Harbom, Strand and Nygard 2008, p. 2). For most cases of spontaneous violent protest, however, this criterion is not met. Nonetheless, acts of violent protest by an ethnic group can escalate into civil war, if they are sustained and reach significant levels of violence (Tilly 2003).

Ethnic Civil War vs. Communal Violence In an ethnic civil war, one of the conflict parties must be the government. For that reason, ethnic civil war is analytically distinct from communal violence, i.e. episodes of conflict between groups in which neither group represents the government. A frequently studied case of this phenomenon are the repeated confrontations between Hindus and Muslims in India (Varshney 2002; Brass 2003; Wilkinson 2004). Communal violence, as Horowitz (2002, p. 1) notes, is an "intense, sudden, though not necessarily unplanned, lethal attack by civilian members of one ethnic group on civilian members of another ethnic group, the victims chosen because of their group membership". The intentions of a group in communal conflicts are very different from those in ethnic civil wars. Whereas the latter are related to broad political issues of the ethnic group at the *national* level, communal violence relates to *local* tensions between groups.

Ethnic Civil War vs. Genocide A genocide is an "organized attempt to annihilate a group thought to have ethnic, racial, or otherwise organic properties" (Straus 2004, p. 88). Genocide can occur with different levels of intensity, ranging from single, localized pogroms, or as large-scale violence with mass participation (Bhavnani 2006). In many cases, genocidal strategies are pursued by government groups in a attempt to create ethnically "pure" territories (Mann 2005). Genocide falls under the category of ethnic conflict, because it occurs between governmental forces and a minority ethnic group. However, even though genocide often occurs in conjunction with civil war, it is not equivalent. This is because violence in civil wars is two-sided, mutually inflicted by both parties upon each other. In contrast, genocidal violence is the systematic killing of civilian populations and therefore does not satisfy the definitional criteria for war.

2.1.1 Ethnic Groups as Conflict Actors

An ethnic civil war pits the government forces against an internal challenger, an ethnic group. What is an ethnic group? Defining ethnicity is notoriously difficult, and there is a substantial lack of agreement in the scholarly community. In fact, the usefulness of the concept has been fundamentally questioned, which led to the suggestion to abandon it altogether (Brubaker and Cooper 2000). Taking a less critical position, this section presents a discussion of ethnicity as a social category in regard to the topic of this dissertation, ethnic conflict.

Groups as Imagined Communities

How can we recognize ethnic groups? One might suggest that ethnic boundaries follow clearly identifiable cultural markers, such as language or religion. Although they do in many cases, an objective definition of this kind fails to take into account the inter-personal meaning that individuals attach to these boundaries (Cederman 2002*b*). What distinguishes an ethnic *group* from an externally imposed ethnic *category* is the *shared belief in a common descent* – in other words, the subjective feeling of belonging (Chandra 2006). Hence, an ethnic group is what Benedict Anderson (1991) – with respect to nations – has aptly called an "imagined community" – a community that attributes its present cohesion to a shared cultural ancestry, whether this corresponds to the historical record or not. Thus, the definition of an ethnic group based on its belief in common descent goes beyond externally observable cultural fault lines, and introduces group boundaries where there are few, if any, cultural differences. For example, the Hutu and Tutsi in Rwanda clearly constitute different ethnic groups. However, both groups speak the same language, and there are no religious differences between them as both groups belong primarily to the Roman Catholic church.

A. D. Smith (1986) spells out the dimensions that define ethnic groups in more detail. All of these dimensions invoke a sense of common group identity, but they do so by referring to different markers or symbols. Most importantly, each group has its *name*. It is virtually unthinkable that a group that considers itself different from others would go under the same name or category as another group. A name helps the group establish its distinctness internally, but also in relation to outsiders. Each group usually has a *common myth of descent*. This is not the real descent according to the historical records, but the putative legend shared by the group members. A common descent is important since it serves as a historical explanation for why the group belongs together. Beyond the common origins that unite group members, Smith mentions the *common history* a group has experienced. The memories of shared experiences in earlier generations links the origins of the groups to present generations. Besides these historical dimensions, groups share a *distinctive culture* that facilitates in-group cohesion and out-group discrimination. Language or religion often coincide with these cultural break lines, but not necessarily so. According to Smith, another observable characteristic of an ethnic group is its attachment to a *homeland*. Again, this attachment does not have to be based on prior settlement of the land, but is represented as such in the group's collective history. Lastly, ethnic groups have a strong sense of *solidarity* among the group members. This explains strong cooperation even between ethnic communities of the same group in different countries.

Smith's definition allows us to distinguish between ethnic boundaries and other social categories. Members of social classes might have a similarly strong feeling of solidarity, but across countries they certainly lack the territorial attachment that ethnic groups have. Moreover, social classes are not characterized by a shared cultural traits, but rather encompass members from different cultural backgrounds with a similar socio-economic status.[1] Ethnic groups do not correspond to religious communities either, because these also lack the relationship to a particular territorial region. On the other hand, a territorially delimited polity alone fails to satisfy the definition of an ethnic group, because it does not necessarily include a shared belief in common descent.

Smith's definition highlights the importance of socially constructed ethnic identity within the group. For the study of ethnic conflict, this necessarily entails the question of whether the concept of an ethnic group provides a useful analytical category. If ethnic groups cannot be presupposed, because they "wax and wane over time" (Brubaker 2004, p. 4), how can we study conflict between them?

[1] An exception is the position advocated by Bourdieu (1984), who argues that class differences provide the basis for the emergence of ethnic groups. Also, Gurr's (1993) "ethno-classes" are characterized by coinciding socio-economic and ethnic cleavages.

Constant vs. Fluid Group Boundaries

The study of ethnic conflict comprises widely divergent theoretical positions with respect to its main concept of interest, ethnicity. Most of these disputes occur along a single dimension, with essentialist or primordialist positions at the one extreme, and constructivist accounts of ethnicity at the other. Primordialists have a static perception of ethnicity. They assume ethnic boundaries to follow cultural markers such as language or religion, and to be largely invariant. Individual membership to a certain ethnic group is determined by kinship, and does not change over time (Geertz 1963). Thus, according to the primordialist perspective, ethnic identity is "hard-wired", both for an individual and also for inter-group boundaries.

At the opposite end of the spectrum, radical constructivist accounts see ethnic identity as a fluid concept that cannot be analyzed independently of the social, economic and political processes in which it is embedded. They argue against the reification of ethnic groups, i.e. the conception of ethnic groups as independent, omnipresent social categories. Chandra summarizes the constructivist stance as follows:

> Those who subscribe to the constructivist approach agree on two basic propositions: First, individuals have multiple, not single ethnic identities; and second, the identity with which they identify varies depending upon some specified causal variable. Changes in the value of these causal variables are likely to lead to changes in individual identifications (Chandra 2001).

What does this mean for the study of ethnic conflict? If ethnic contestation or conflict is one of the causal variables that affects individuals' identification, it is difficult to study the impact of ethnic variables on conflict, since the two are so highly interdependent. This is the reason why some authors have argued against the widespread "groupism" in the social sciences, i.e. the tendency to use ethnic groups as the unit of analysis (Brubaker 2004). However, instead of dismissing the concept of a group altogether, various scholars have advocated an intermediate position, starting with the pioneering work of Fredrik Barth (1969). Instead of fixed group membership based on cultural traits as advocated by primordialists, Barth emphasizes that it is only the boundaries that are static, but not group membership. In his own words,

> boundaries persist despite a flow of personnel across them. [..] [D]iscrete categories are maintained *despite* changing participation and membership in the course of individual life histories. (Barth 1969, p. 9-10, emphasis original)

Whereas Barth emphasizes the persistence of boundaries in general, other scholars argue that the "flow of personnel across them", as Barth calls it, is limited. In her convincing attempt

to provide a definition of ethnicity, Chandra (2006) sees ethnic identities as social categories in which membership is determined by (real or putative) descent. Based on this assumption, she derives two attributes of ethnic categories: constrained change and visibility. The former refers to the fact that it is difficult to change descent-based attributes that define ethnic groups, such as common skin color, native language, or place of birth. While the visibility of these attributes varies, the membership they define might be fairly sticky as compared to other social categories. Because these group memberships are fairly static, we need not dismiss the concept of an ethnic group altogether. Group boundaries persist, and constitute relevant societal fault lines. In the remainder of this dissertation, I therefore take a perspective as advocated by Van Evera (2001) and treat the existence of groups as exogenous to conflict, while fully acknowledging that this might not be applicable when we deal with ethnicity in the context of other societal phenomena.

The Role of Groups in Conflict

The contemporary literature on ethnic conflict tends to treat ethnic groups as the main conflict actors. For example, the title of Donald Horowitz's well-known book, *Ethnic Groups in Conflict* (1985), seems to suggest that ethnic conflicts are mass events, with "all-against-all". This approach has come under some criticism, advanced most importantly by Mueller (2000). Mueller portrays the ethnic conflicts in former Yugoslavia and Rwanda as acts of violence perpetrated by "well-armed thugs and bullies" (p. 43) that were encouraged and paid for by political authorities seeking personal advantages. Clearly, this puts the motivations for conflict far away from the deep-rooted grievances and hatred that the literature traditionally holds responsible for violence. Similarly, Gagnon (2004) argues against the conception of ethnic conflict as war of groups against groups. Rather, he holds that political entrepreneurs are responsible for the outbreak of violence, as these leaders attempt to fend off mobilization attempts by the opposition that threaten their political survival. Most important for our purpose, however, is the fact that the participation of the broad population is not a necessary condition for conflict to occur.

Clearly, only in rare cases do entire ethnic groups actively join in hostile acts. However, there are two reasons why I argue against the dismissal of ethnic groups as an analytical category. First, we need to be clear that the active application of violence against members of another group is only the most extreme form along a continuous repertoire of actions. Above, I have defined the phenomenon of ethnic civil war as a conflict along ethnic lines regarding claims relating to ethnic groups. What is assumed to be driving these conflicts is a collective grievance at the group level, a shared political goal that leads a subset of the group to take arms. Consequently, in order to find out why conflict occurs, we need to find out how these collective grievances come into existence in the first place. This strategy follows the approach in the

social movements literature, as outlined by Tilly (1978). Tilly distinguishes three analytical perspectives in the study of revolutionary collective action: groups, events and movements. For our purpose, the distinction between groups and movements is helpful. A movement is a set of individuals that take (possibly violent) action to achieve some goal. However, this goal is defined with respect to a broader group of common structure and shared beliefs. In other words, the movement is a subset of the group that takes manifest action to further the group's interests. The group–movement distinction also applies to ethnic conflict, where the ethnic group provides the collective motivation for which certain organizations choose the most extreme form of action: the application of violence. The production of violence within these organizations is an important question, but beyond the focus of this dissertation. However, efforts are underway to disaggregate ethnic groups at the level of the organizations representing them. For example, the *Minorities at Risk Organizational Behavior* dataset (MAROB, Minorities at Risk Project 2006) aims to create a list of organizations acting on behalf of ethnic groups, including variables that measure the degree of violence these organizations choose to apply.

There is, however, a second reason why groups matter as units of analysis. As mentioned above, Mueller and Gagnon reject the group level, because violence is perpetrated by small subsets of the full population, without mass mobilization. However, what about the targets of violence? Various accounts of the Bosnian war show a strategic targeting of enemy group population in these conflicts (e.g. Mann 2005; Burg and Shoup 1999). Clearly, on the target side, violence was not restricted to small subgroups, and ethnic membership was key for the perpetrators in selecting their targets. Consequently, the question remains why groups – and not individual extremists – develop strong antagonism against each other, even though not all their members join in the fighting.

Having discussed the definition of ethnic civil war and groups as the actors in these conflicts, I now turn to the existing theoretical explanations for ethnic conflict.

2.2 Why People Fight: Existing Approaches

Ethnic civil war has received scholarly attention from a wide range of disciplines beyond political science. Psychologists, sociologists, anthropologists and economists alike have been trying to explain the phenomenon of violent conflict between groups. Not surprisingly, there is considerable variation with respect to theoretical approaches. This section gives an overview of the existing explanations, with a focus on the political science literature on ethnic war. Explanations for ethnic conflict can roughly be put in three categories. The first category analyzes conflict as occurring between groups but not necessarily inside the context of a state. Here,

conflict is either taken as an inter-cultural phenomenon where states do not matter (as in ancient hatred theories), or conflict takes place under the condition of a temporary absence of a state. I label this category "symmetric conflict", since the groups are treated as equal actors. In contrast, "asymmetric" ethnic conflicts of the second category occur between groups with a different (political or economic) status within in a given state. As a third category, the recent political economy literature on civil war attributes little explanatory value to ethnicity.

2.2.1 Symmetric Conflict: Groups vs. Groups

Explanations of the first category examine conflict outside the context of a state.

Ancient Hatred

Above, I have introduced the essentialist perspective of ethnicity. Essentialists take the existence of group boundaries as given and see group membership as largely invariant. However, this alone does not explain why conflict arises between groups. Ancient hatred explanations of ethnic conflict see a persistent enmity between groups as the root cause of inter-ethnic hostilities (see e.g. Geertz 1963). Culturally different groups are assumed to live in a constant state of enmity such that violence will result whenever external conditions allow it. Huntington's *Clash of Civilizations* is an example of such an essentialist explanation for conflict. Huntington (1993, 1996) argues that the end of the Cold War would give rise to a new era of conflict between clusters of "civilizations", whose boundaries largely follow religious lines. Among the most belligerent of these clusters, according to Huntington, is Islam. Even though ancient hatred explanations no longer receive much attention in the scientific community, they continue to be present in the media discourse (Newman 1991). Robert Kaplan's book *Balkan Ghosts* (1993) is a prominent example in which he states that the reason for the Balkan conflict lies in the region's "belligerent culture".

The Ethnic Security Dilemma

The *security dilemma* approach to ethnic conflict grew out of the realist school in international relations (Posen 1993*a*,*b*). Realism in international relations rests on the basic assumptions that states are the main actors in the international system, and that their primary concern is their own security. The system itself is anarchic, so there is no central authority that is able to guarantee the integrity of states. As a result, states are responsible for their own security and therefore take preemptive military measures. Posen applies this thinking to an intrastate setting with groups as the main actors. He reasons that during the collapse of multi-ethnic states, groups live in a

state of temporal anarchy because the federal government that used to protect their interests is declining in power. Groups therefore have incentives for military action against other groups.

Rational Choice Explanations for Ethnic Conflict

Different attempts have been made in the rational choice literature to explain ethnic conflict as the result of strategically interdependent actions by ethnic groups. This literature departs from the assumption that groups are "unitary" actors, that act in unison with the goal of maximizing some expected outcome along the political or economic dimension. There are a broad variety of works on ethnic conflict employing a rational choice approach, a comprehensive overview of which is given by Fearon (2006). Here, I focus on two examples from this literature, one of which examines conflict between groups facing the decision whether to enter a common state, and the other one explaining ethnic conflict as the result of the failure to share the spoils in an existing state. A game-theoretic model by Fearon (1998) relates the emergence of ethnic violence to a commitment problem. Similar to the security dilemma theory of ethnic conflict, Fearon assumes ethnic conflict to occur during or shortly after the breakup of multi-ethnic empires. A group facing the decision whether to enter a new state together with a majority group will refrain from doing so because the majority group may not keep its promises once the state has been created. The minority might thus opt for secession before the majority can eventually consolidate its power, which would make secession impossible in the future.

Caselli and Coleman (2006) advance a model where the struggle of groups for a society's wealth-producing assets can lead to conflict along ethnic lines. At the core of their model is the assumption that the assets in a society are controlled by a winning coalition of agents. However, once this coalition has assumed control of the assets, it is beneficial for non-members to infiltrate the coalition in order to get a share of the spoils, which is of course unfavorable for the coalition members, since their share of the benefits would decrease. The authors argue that ethnicity serves as a marker to identify these potential spoilers. A coalition group defined along ethnic lines is thus more likely to fight for the control of a country's assets, because the potential benefits are higher than those of a coalition without ethnic markers. As a result, conflict should be more likely in ethnically heterogenous countries.

2.2.2 Asymmetric Ethnic Conflict: Groups vs. States

In contrast to the approaches presented in the previous section, the majority of theoretical approaches to ethnic conflict examines inter-group relationships in the context of a state, where political or economic discrepancies between groups in a state can lead to inter-group conflict.

Relative Deprivation and Ranked Groups

Theories of relative deprivation, advanced by Gurr (1968, 1970), claim that a group's violent reaction is caused by the perceived discrepancy between what the group thinks it is entitled to as members of a state, and what they get in reality. This discrepancy can be measured along an economic dimension, for example if the group is systematically excluded from high-income positions. Alternatively, relative deprivation can also occur politically, for example by not granting group members access to political elite positions, or otherwise not allowing the group to participate in politics. This perceived deprivation, according to Gurr, leads to a feeling of collective frustration of the group. Gurr draws on the "frustration–aggression" hypothesis from the psychological literature to link perceived deprivation to collective violence (Dollard et al. 1939).

In principle, the effect of relative deprivation as a motivation for conflict is not restricted to ethnic groups only. However, it is most likely to be effective if ethnic group boundaries coincide with socio-economic divisions. This is what Horowitz (1985) calls a "ranked" system of ethnic groups. In a ranked system, groups differ with respect to an observable status – for example income or political position. However, unlike Gurr, Horowitz's account does not assume ranked systems to be more conflict-prone in general. Whereas he acknowledges the fact that a ranked system certainly has sufficient potential for violent escalation, the asymmetry in the status of groups can also reinforce the more powerful groups in repressing the subordinate ones. This is why an unranked system in which the social structure of the groups spans different levels of the socio-economic hierarchy can be much more dangerous, since the conflict occurs between opponent groups of equal strength.

Elite Manipulation

Theories of elite manipulation (Gagnon 1995, 2004; de Figueiredo and Weingast 1999) see ethnic conflict as the result of intra-group dynamics between group elites and the masses. The ruling group elite, when afraid of losing power in the domestic arena, will play up ethnic differences and actively antagonize outside groups in order to mobilize intra-group support for their political survival. As Gagnon puts it,

> by constructing individual interest in terms of the threat to the group, endangered elites can fend off domestic challengers who seek to mobilize the population against the status quo, and can better position themselves to deal with future challenges (Gagnon 1995, p. 132).

Elite manipulation theories cast ethnic conflict largely as an elite enterprise. The active mobilization by belligerent leaders seems to be sufficient to engage a group in conflict. However, it is difficult to disentangle the causal impact of leaders from the group grievances they choose to mobilize around.

Ethno-nationalism

At the root of ethno-nationalist conflict is the nationalist sentiment of ethnic groups. In the words of Ernest Gellner, "nationalism is a political principle, which holds that the political and the national unit should be congruent" (Gellner 1983, p. 1). Ethnic conflict arises out of this principle if an ethnic group that is presently part of a state develops claims to be a "nation" and demands independent statehood. To be sure, not all ethnic groups are nations. As we have seen above, the definition of an ethnic group does not include a political dimension (Cederman 2002*b*). However, once an ethnic group becomes a nation, ethnic boundaries become political boundaries, and conflict results from the attempt to match them.

According to this perspective, nationalist sentiment is the root cause of conflict. However, why does nationalist sentiment emerge in the first place? The literature on nationalism gives different answers to this question. Gellner (1983) attributes the rise of nationalism to the increased importance of vertical cultural integration in industrialized societies. Pre-modern societies consisted of insular peasant communities ruled by an upper stratum of aristocratic and cleric elites. In the age of industrialization, however, the dependence of the upper and lower classes on each other makes a vertical integration of the classes necessary. This was provided by a cultural integration under the umbrella of nations, homogenizing them internally and setting them apart from others. Anderson (1991) highlights the impact of state schooling and a common language in creating a national community. Only when people speak the same language does it become possible to imagine a community of common descent.

Nationalist ambitions by ethnic groups can take a variety of forms, and limited space prevents me from discussing them in detail. Probably the most common form is *peripheral* nationalism, where an ethnic group that is part of another state attempts to secede in order to create its own state (Hechter 2000*a*). However, in many cases these secession attempts can be spurred if the ethnic group has a kin group living nearby, and tries to unite with this group in a new state. This is what Hechter calls *unification* nationalism. In contrast, if the kin group dominates a nearby state and the aim of a seceding group is to join this already existing state, we talk about *irredentist* nationalism (Hechter 2000*a*; Weiner 1971).

2.2.3 Ethnicity in the Recent Civil War Literature

The recent years have seen a surge of interest in the study of civil war. Mostly conducted by political economists, much research focuses on the question of whether economic conditions or group motivations are better explanations for the onset of conflict. Essentially, this literature challenges the existing theoretical explanations as described above – not by simply proposing an alternative approach that links ethnic groups and conflict, but by questioning the relevance of ethnic boundaries altogether. Rather than ethnic "grievances", these authors postulate that the causes of civil war are rooted in individual "greed" for material gains.

The recent literature has identified various links between poverty and civil war. In the following, I focus only on two contributions from this literature, as they are the most influential ones. The first causal explanation was proposed by Collier and Hoeffler (2004) in their paper "Greed and Grievance in Civil War", a title that summarizes the clash of different theoretical schools. According to Collier and Hoeffler, people fight in order to improve their economic position. In countries with low income and a poor educational system, there are few prospects of earning an ordinary living. Rebellion thus becomes an attractive alternative, a way to make profit where all else fails. The second mechanism linking economic performance and civil war stems from Fearon and Laitin (2003), who attribute the outbreak of civil war in poor countries to a weak state. Poor states, they argue, have fewer capabilities of repressing rebellion and thus have a higher conflict risk. Both theoretical links obviously do not require collective grievances at the level of ethnic groups as a necessary condition for conflict, unlike the other approaches outlined above. Collier and Hoeffler's opportunity-cost argument applies to any citizen of a poor state. For poor people, rebellion is a feasible source of income, so any poor person can choose that option regardless of their ethnicity. A strong state, as in Fearon and Laitin's explanation, would suppress any public unrest – not necessarily at the group level – that would otherwise result in conflict.

The proposed causal links are tested in large-N empirical studies at the country level, using aggregate indicators of ethnicity – mostly the index of *ethno-linguistic fractionalization* (ELF), which gives the probability that two randomly selected individuals will be from different ethnic groups. Empirical findings support the proposed causal links: Economic performance is strongly and consistently related to civil war risk. Most importantly, however, economic performance trumps the ethnic indicators, for which no such relationship was discovered. This led to the conclusion that ethnicity is irrelevant as an explanation for civil war, once material factors have been controlled for. However, we need to keep in mind that this conclusion is based on a single aggregate indicator of ethnicity, ELF. What if ethnic diversity as measured by ELF does not capture the true processes linking ethnic groups to conflict? The critiques by Cederman and

Girardin (2007) and Buhaug, Cederman and Rød (2008) show that a different operationalization of ethnicity does in fact reveal a significant impact of ethnicity on conflict, contrary to the conclusions of the political economy literature on conflict.

2.2.4 Summary: Existing Approaches and Open Questions

In summary, we can distinguish between two broad bodies of literature. Before the end of the Cold War a rich literature on ethnicity and conflict emerged, emphasizing different motives groups might have to resort to arms: Scholars of nationalism attribute ethnic conflict to the rise of a national awareness and the resulting claims for statehood. Ancient hatred theories assume an innate enmity between groups as the cause of violence. Groups can fight because they fear for their security, as the result of mobilization by belligerent leaders, or because they feel disadvantaged in comparison with other groups. Whereas this literature is strong on the motivation of groups for conflict, it does not answer the question of how this collective motivation develops into violent conflict. The above definition of ethnic civil war should have made clear that we are not talking about minor instances of spontaneous violence. Instead, we see enduring episodes of large-scale violence, which, by definition, require the involvement of a large number of people.

Whereas the earlier literature focused almost exclusively on ethnic grievances as motivation for conflict, the more recent body of literature written after the end of the Cold War favors opportunities as explanation. If rebel recruitment is facilitated by a poor economy, or if state repression is weak, violent conflict will arise. According to this perspective, motivations for conflict can be presupposed and do not deserve further scrutiny – instead, opportunities drive civil wars. Similar to the motivational literature, however, the recent literature can also be criticized for its lack of focus on causal mechanisms. By exclusively relying on country-level indicators, these studies cannot give us much indication as to how large groups mobilize for violence. Sambanis (2004*a*) summarizes this shortcoming as the result of an over-aggregated research design, with a focus on *factors* rather than *actors*.

In order to remedy the shortcomings of the literature, recent work has often taken a disaggregated level of study, trying to study civil war and the state it occurs in at a finer level of detail. This dissertation is part of this research tradition. The next section describes the disaggregation approach in more detail.

2.3 Disaggregated Studies of Civil War

Generally speaking, disaggregation is the attempt to theorize, measure and test at a more precise resolution by breaking up the composite object(s) of study into their constituent parts, and

analyze how these parts work together in order to create the outcome at the aggregate level. This definition suggests that for "disaggregation" to be possible, there must be an aggregated concept, something whose smaller pieces can be examined. Most – but not all – studies in the disaggregation tradition take country-level aggregations as their reference point and disaggregate them to the sub-national level. The main motivation behind disaggregation is to get closer to the micro-level processes that drive civil war. Simply observing macro-level relationships such as the correlation between economic performance and civil war leaves us unsure as to how macro-conditions bring about the macro-phenomenon of conflict. For a good explanation, we must examine *micro-mechanisms*, i.e. how these macro-conditions affect the interactions of individuals, and how these in turn create the conflict outcome (Hedström and Swedberg 1998; Coleman 1990).

Scholars studying mechanisms usually advocate qualitative approaches such as process-tracing as their empirical method of choice (Checkel 2008; George and Bennett 2004). In a process-tracing analysis, the hypothesized process is carefully scrutinized by examining the interactions between the intermediate causal agents involved. For example, group mobilization for conflict could be studied on a past case of ethnic conflict, carefully tracing how unrest spreads through the population, eventually reaching the level of full-fledged war. Process-tracing requires a detailed case analysis of the actors and events that are causally prior to the phenomenon in question. Typically, only a few cases can be examined at such a high level of detail. Moreover, in process-tracing case studies it is difficult to establish causality: Would the outcome be the same if some event A had not occurred, or if some actor B had reacted differently? The advantage of process-tracing is that the analysis is tightly linked to the causal agents and their actions. However, this advantage exists at the expense of generality. For example, case studies on the civil war in El Salvador (Wood 2003) or Greece (Kalyvas 2006) cannot tell us whether the proposed relationships can be generalized to other cases beyond those studied.

What I refer to as the disaggregated research tradition differs from a purely qualitative case study approach. Disaggregated studies still use a large-N design, by employing much higher resolution data. Different strategies exist for disaggregation. One possible categorization distinguishes between *conceptual*, *spatial* and *temporal* disaggregation. Conceptual disaggregation is a theoretical refinement of an existing argument by examining the constituting parts of a theoretical concept. For example, Goemans, Gleditsch and Chiozza (2009) study the different ways political leaders enter and leave office, and how this effects regime stability. Cederman and Girardin (2007) propose an alternative conception of ethnic conflict by distinguishing between politically included and excluded groups, rather than explaining conflict as the result of ethnic heterogeneity at the country level (Fearon and Laitin 2003). Spatially disaggregated studies

consider the geographic variation in a particular variable of interest. A prominent example is the study by Buhaug and Gates (2002) on the spatial extent of conflict within a country. Further work has pushed spatial disaggregation to even finer levels of resolution, for example by studying the spatial correlates of African civil wars at a 100x100 km resolution (Buhaug and Rød 2006). Finally, temporal disaggregation singles out the particular events of a conflict, rather than observing conflict at the country-year level. For example, Weidmann and Ward (2008) examine the dynamics of conflict at a monthly resolution.

An alternative categorization of disaggregated research distinguishes between the disaggregation of the *actors* involved in a conflict, the *context* that shapes their interactions, and *conflict* as the dependent variable. The disaggregation of actors has been achieved by e.g. looking at the internal organization of a rebel group instead of treating it as a unitary actor (Gates 2002). Many studies do what I label context disaggregation: For example, Hegre and Raleigh (2006) test the impact of population on conflict by employing population estimates at a high spatial resolution. The same study also disaggregates conflict spatially: Using data from the ACLED dataset (Raleigh and Hegre 2005), the dependent variable in the study is the occurrence of conflict *events* rather than civil conflict.

Disaggregation does not solve all the shortcomings of the current quantitative literature on conflict. First, since technological advances in data collection and processing allow for an almost unlimited refinement in the level of analysis, researchers run the risk of "over-disaggregation". A higher empirical resolution will not necessarily enable us to better capture the mechanisms we are interested in. Rather, civil wars are driven by different mechanisms, each of which operate at a different level of analysis (Weidmann 2009). A second criticism regarding disaggregation is the absence of actors. Despite the promise of getting closer to the mechanisms, many studies still produce correlations of factors, although at a much finer level. In order to further our understanding, however, we need to explain how the actions of individuals generate the observable conflict outcome. Taking these two caveats seriously, the next section shows how the study of group settlement patterns can bring us closer to the processes of intra-group mobilization and the organization of violence.

2.4 Groups, Geography and Conflict Mechanisms

Over the past decades, the social sciences have started to take geographic space more seriously. Whereas before, space used to be outside the scope of the analysis, the "spatial turn" brought with it an increased awareness of the importance of geography for social behavior. This perspective is for example expressed in the works of the sociologist Anthony Giddens who argues

that human behavior is always embedded in a spatial and temporal context, and as a result, also structured by this context. At the same time, however, this behavior produces and reproduces this context (Giddens 1984).

The role of space in the social sciences can broadly be distinguished along two perspectives: First, the *physical properties* of space as determinants of social behavior, and second, space as providing the *interaction opportunities* for social agents. An example of the former perspective is André Siegfried's explanation for regional variation in French voting patterns (Siegfried 1975).[2] Siegfried claims that the geology of a particular region determines the voting pattern (on a left-right dimension). He draws a link from the soil type to land use, which again has an impact on property relations in a region and thus on voting behavior. Similarly, in the field of civil war research, some scholars posit a relationship between certain physical properties of geography and conflict. For example, Lujala, Gleditsch and Gilmore (2005) propose a link between the occurrence of diamonds as a lootable resource and the likelihood that a region will be affected by violence.

The second perspective sees geography as a topology that determines the interaction opportunities of social agents. In other words, in a space where actors are assigned to particular locations, the space determines to a certain degree who can interact with whom. Here, space does not have an immediate causal effect on the outcome of a particular social interaction, but defines the set of interactions that effectively occur. We can broadly distinguish between interactions at different geographic levels. At the macro-level, the geography of the international system determines the interaction of states. Below the state, at the meso-level, regional groups and organizations constitute the first layer of subnational entities. These groups in turn consist of individuals with their particular geographic location at the micro-level.

This dissertation follows the second perspective and studies how geographic interaction opportunity is linked to conflict. This is not a new theme in conflict research more generally. In the study of international conflict, it has been shown that the number of interstate conflicts increases with geographic proximity between the opponents (Richardson 1961; Wesley 1962). According to this *interaction perspective* (Hensel 2000), states interact more with their close neighbors, which in turn increases the likelihood of violent dispute. However, here I assume a much more narrow focus. I examine interaction opportunity at the micro- and meso-level and how it affects the risk of ethnic conflict. Micro-level interactions between group members will be examined in detail in Chapters 4 and 5. Here, the group settlement pattern is assumed to determine the interaction opportunities of group members for conflict mobilization – in other words, geography can make a group more prone to conflict by facilitating intra-group mobiliza-

[2]Thanks to Heiri Leuthold for alerting me to this example.

tion. Chapter 6 focuses on the meso-level. I study the interaction opportunities between groups defined by their settlement patterns, and how this affects the location of violence. The following two sections present the theoretical background for geography as a determinant of interaction opportunity, both at the micro- and meso-level.

2.4.1 Micro-level: Interaction Opportunities within Groups and the Risk of Conflict

The relationship between group territory and conflict has long been mentioned by scholars, but it was not until recently that it began to receive more systematic coverage. Early accounts include Bates (1974) who holds that the proximity of group territory to the developed centers of a country strengthens the influence of modernization and might therefore give rise to an ethnically stratified society with a high potential for conflict. However, Bates's argument applies to the relative locations of the different group territories to each other. The territory itself, according to Bates, is always geographically concentrated: "There is no denying that the members of an ethnic group tend to cluster in space" (Bates 1974, p. 464). Here, territorial concentration is the degree to which an ethnic group lives clustered together in a country. A perfectly concentrated group occupies a single, contiguous region, whereas a dispersed group lives in smaller settlement patches distributed across the country. The study of group concentration received more attention with the creation of Gurr's *Minorities at Risk* dataset (Gurr 1993). The purpose of the MAR dataset was to inductively learn more about the various driving forces behind minority conflict around the world. Gurr distinguishes between two major types of groups – *national* and *minority* peoples, the former being "regionally concentrated". However, Gurr does not go beyond a descriptive assessment of group concentration and conflict. A more systematic assessment is performed by Toft (2002, 2003). According to Toft, one of the reasons why concentrated groups face a higher conflict risk is that group concentration facilitates collective action and is thus related to a higher conflict propensity. This corresponds to the accounts of other authors who have argued in along similar lines (Lichbach 1995; Hardin 1995).

Even though the effect of geographic concentration on social interaction has been mentioned in the literature, little attention has been given to its detailed specification and testing. How exactly does geography affect group mobilization for collective action? In order to answer this question, I draw on existing work on social movements. Research on social movements has generated a large body of literature on mobilization mechanisms of groups; however, to date this literature has not had much intersection with the classic literature on ethnic conflict introduced above. Clearly, the focus of the social movements literature is on the practical aspects of collective unrest – how does a large number of people that share a certain grievance get together

and take action? In focusing on motivations for ethnic conflict, the existing literature has largely omitted this question, which is why a closer look into social movements research might prove helpful.

In short, my argument is as follows: Group settlement patterns determine the interaction opportunities of individuals. If close and frequent interaction is possible, mobilization can easily spread through the group population such that collective action becomes likely. This argument rests on two assumptions. First, individual behavior has a strong social component. In other words, the decision an individual makes is strongly influenced by her peers. In the case of ethnic conflict, this means that the decision whether to participate in collective action is strongly dependent on the influence of the ethnic kin. Second, I assume that the influence of ethnic peers is contingent on spatial proximity, such that more proximate individuals will have a higher impact on each other's actions. I will now discuss these assumptions in turn.

Individual Behavior is Socially Influenced

The credo of the structuralist approach in the social sciences is that the behavior of individuals cannot be studied in isolation, as it depends crucially on the social network – the *structure* – the individual is embedded in. As Zuckerman puts it: "When persons make decisions, they take into account the cues, knowledge, values, and expectations of their spouse, parents, children, friends, work-mates, and others around them – those who matter in their lives" (2005, p. 1). This idea is not new. In fact, it can be traced back to the work of the sociologist Lazarsfeld and his collaborators in the 1950s (Lazarsfeld 1968). Social peer influence affects all sorts of decisions, but most importantly for our purpose, it affects whether an individual participates in a social movement (Pinard 1968; Tilly 1978). How exactly do interpersonal links affect individual decisions? Diani (2003, p. 8) identifies three types of influences: interpersonal links as a source of (i) socialization, (ii) information, and (iii) social pressure. The following paragraphs discuss these types of social influence.

Socialization Shared goals and values is a necessary condition for collective action to arise. Only if the group members are aware of their shared grievances and motives is it possible for collective action to be taken to attain a common goal. Social interaction is the process by which these shared feelings are developed among a population. Hardin mentions the example of Marx's theory of revolution

> which requires the development of a class consciousness before there can be class-oriented action. Workers in a factory share so much time together that they begin to

> understand their common fate much better, not least because each can benefit from the insights of all (Hardin 1995, p. 55).

In the case of ethnic groups, does social interaction matter for the emergence and salience of ethnic boundaries? Certainly not, if we subscribe to an essentialist perspective on ethnic groups. However, using the insights from constructivist theory I have presented above, it is clear that social interaction between group members is a requirement for the ethnic group to be internally recognized as such. Only if group members have developed a shared feeling of belonging by mutual interaction will ethnic boundaries become internally salient. Moreover, in the case of ethnic groups, the institutional level will add to the socialization effect. In many cases, ethnic groups enjoy a certain degree of regional autonomy. This entails some material benefits for the group, such as tax revenue (Hechter and Okamoto 2001), but also allows the group to retain a certain degree of distinctiveness. In these settings, the socialization of group members to a common group identity is likely to be strong. Above, I have mentioned Anderson's (1991) theory of the emergence of nationalism, which rests on the assumption that a feeling of belonging develops when there is a common language and state schooling. This effect is likely to be strong in the case of ethnic groups, since the control of the regional education system is often held by the group.

Information Interpersonal networks provide the basis for the flow of information between group members. This information can be about current and planned activities, or simply the current level of participation in collective action. The idea that individuals base their participation decision on the current strength of a movement is central to Granovetter's threshold model of collective action (Granovetter 1978). According to this model, an individual requires a certain number of participants (thus *threshold* model) before they themselves join in. This is so because the current level of participation allows the individual to assess the potential benefits of joining: If current participation is low, the risk of being arrested is high, and the overall likelihood of success is low. On the other hand, if participation is high, collective action is likely to succeed, so individuals will be more willing to join.

The assumption that the current level of support for a particular collective cause influences individual decisions is part of various explanations for collective action. Just to name one other example, Kuran (1989, 1991) models the decision to join a rebellion as a result of two considerations: first, the utility an individual derives from acting according to her personal beliefs, and second, the reputational utility from being known as a supporter. Reputational utility is again dependent on the current level of participation. If many people join, it becomes increasingly difficult not to join as well. The current level of support for a movement, however,

is often not revealed publicly. Rather, an individual will most likely obtain this information from friends and colleagues. This is why in the context of mobilization for collective action, the spread of information along immediate social contacts is crucial in influencing people's participation.

Social Pressure Besides the socialization and information aspects of peer influence, peers can simply exert social pressure to make people participate. In its weakest form, social pressure amounts to encouraging prospective members (Diani 2003). However, social pressure can also take the form of physical coercion with the goal of forcing people to join. We must not forget that collective action up to the level of large-scale physical violence might bring significant risks (e.g. of being arrested or even killed). Collective action theory tells us that under these circumstances, it is a better option for an individual not to join and to let others do the work, while still enjoying the benefits of the collective effort (Olson 1965). Therefore, groups monitor participation carefully and administer punishments in the case of non-participation. This peer pressure is important and has been found to have a much higher impact than external incentives (Moore 1995; Fearon and Laitin 1996). In other words, in addition to the benefits that can be achieved by the collective action of a group, there are also potential costs of non-participation to be taken into account, which stem from the social influence of peers.

Spatial Proximity Increases Influence

In the previous section, I have argued that the decision for participation in collective action depends crucially on the influence of peers. The purpose of this section is to link geography to peer influence – essentially, I argue that more proximate peers have a higher degree of influence. This is, in a sense, the First Law of Geography – that "everything is related to everything else, but near things are more related than distant things" (Tobler 1970, p. 236) – applied to the social context.[3] Again, I examine the spatial dimension of peer influence under the three categories introduced previously: socialization, information and social pressure.

The idea that lies at the core of this relationship is that even though modern communications technology seems to overcome the limitations that geographic distance imposes on interpersonal contacts, most social networks correlate highly with spatial proximity. In other words, people are most likely to have strong social ties to other people with whom they can regularly interact on a personal level, because they are spatially proximate (Larsen, Urry and Axhausen 2006). Thus, spatial proximity increases social influence by facilitating interaction.

[3] See Miller (2004) for a discussion of Tobler's first law in the context of human geography.

Socialization The effect of spatial proximity on peer influence has been studied in social psychology. Latané's *Theory of Social Impact* (Latané 1981; Nowak, Szamrej and Latané 1990) isolates three dimensions that affect the impact of the group on an individual's beliefs and attitudes: (i) the *strength* of the group, i.e. the importance the group has for the individual; (ii) the size of the group, with the assumption that a larger group will typically exert a higher influence; and (iii) the immediacy of the group, i.e.its proximity in space and time. The immediacy dimension entails that people that are closer in space are more likely to have an impact on each other's beliefs and values.

The role of geographic proximity for the establishment of a shared group identity is also present in the work of Hardin. Above, I quoted his example from Marx's theory of class conflict, where he holds that frequent interactions between factory workers are a pre-condition for the development of a class consciousness. The example continues by addressing the impact of geographic location: "Peasants scattered across the countryside cannot spend enough time together to gain a comparable sense of class identity" (Hardin 1995, p. 55). Lichbach (1995, p. 158-159) makes a similar point. In studying the "rebel's dilemma", i.e. the difficulties for rebels to gather a group for violent collective action, he mentions group concentration as one of the factors facilitating this enterprise. Spatial concentration, according to Lichbach, contributes to frequent interactions between dissidents, which brings about a "cognitive proximity" in terms of shared grievances and goals.

Information Spatial concentration also improves the flow of information between group members, which in turn facilitates group mobilization. Above, I have mentioned the importance of information about the current level of participation in convincing others to join. Essentially, prospective members of a social movement need signals or cues that indicate they will not be fighting alone. On the one hand, overt signals can be prevented by the authorities, so group members must rely on informal channels to disseminate these cues. Concentration helps to solve the coordination problem of collective action, as group members can interact more closely to send these signals (Lichbach 1995; Hardin 1995). Furthermore, the emergence of visible action can serve as a signal that the movement is growing, which will again lead to increased rates of participation (see e.g. Lohmann 1993). This signal can also have a spatial effect in that it is perceived more strongly in geographically close locations. In fact, episodes of growing civil unrest have been labeled "forest fires" to underscore the spatial pattern of diffusion. For example, Biggs (2005) finds empirical evidence consistent with a spatial diffusion model of of strikes in Paris and Chicago. This spatial pattern also seems to exist for the diffusion of protests in the American South in the 1960s (Andrews 2006).

Social Pressure Spatial proximity between group members contributes to mobilization because social pressure for participation is likely to be higher between close individuals. This is due to two factors: First, immediate contact between individuals makes it easier to detect non-participation. In order to make collective action feasible, deviation from the norm of collective action needs to be identified and punished. This is much more effective under close social interaction in the dissident community. Lichbach (1995, p. 214) highlights the importance of mutual monitoring between group members as a way to detect compliance. Second, once identified, spatial proximity makes it possible to administer punishment to those who have defected. Lichbach (1995, p. 242) mentions three principal techniques of punishing non-contributors. First, the names of non-contributors are published to provoke feelings of shame, second, group members are ostracized by the movement. The most powerful technique, however, is the application of violence against defectors in order to coerce participation, but also to deter defection by others.

Summary: Settlement Patterns and Conflict Mobilization

In the previous section, I made an attempt to anchor the relationship between group concentration and conflict in the social movements literature. I argued that the decision of whether to participate in an ethnic uprising is strongly influenced by peers. If the group geography favors a close interaction between group members, group mobilization is feasible. On the other hand, if geography largely isolates individuals, the emergence of collective action on behalf of the group is unlikely. This logic is based on the geographic opportunities for successful group mobilization, and leaves out the motivation for fighting as emphasized by much of the literature on ethnic conflict introduced above. In an empirical test of this argument in Chapters 4 and 5, it will therefore be necessary to establish group motivation and test the impact of settlement patterns on a set of groups that would have a reason to fight.

2.4.2 Meso-level: Interaction Opportunities between Groups and the Location of Violence

Whereas the previous section examined the relationship between settlement patterns and conflict under an intra-group mobilization perspective, I now turn to settlement patterns as a determinant of conflict between groups. This perspective differs from the previous one with regard to two aspects. First, unlike in the previous part where my unit of analysis was a single group, here I focus on settlement patterns and conflict in a dyadic setting, i.e. how group geographies *in relation to each other* affect conflict. Second, the emphasis of this section is on the *location*

of violence rather than the occurrence of conflict in general. I argue that settlement patterns create strategic incentives for groups to apply violence at particular locations. In other words, settlement patterns do not provide an explanation of whether conflict starts in the first place, but rather where violence is applied once an ethnic conflict has started.

Dyadic Group Geography and Conflict

The argument that geography influences conflict by determining interaction opportunities between actors can be found in different variations in the literature. As mentioned above, studies on international conflict reveal that there is more conflict between close states (Hensel 2000). Rather than focusing on state actors, Huntington (1993, 1996) uses the interaction theory to explain what we calls conflict between "civilizations". Huntington divides the world into eight civilizations: the "Western, Confucian, Japanese, Islamic, Hindu, Slavic-Orthodox, Latin American and possibly African civilization[s]" (Huntington 1993, p. 25). He hypothesizes that the conflicts of the future will occur between groups and nations that belong to different civilizations. On a geographic dimension, the fault lines between the clusters of civilizations will be where we see violent conflict erupting. For example, according to Huntington the wars in former Yugoslavia occurred where they did because the region is located at the fault lines between the Orthodox, Muslim and Western civilizations. However, Huntington's thesis is not supported by empirical evidence, as different analyses have shown (Fox 2002; Gartzke and Gleditsch 2006).

Security dilemma explanations of ethnic conflict also emphasize the role of dyadic group geography, i.e. of group settlement patterns in relation to each other. Posen (1993*a*,*b*) argues that the geographic distribution of groups can create situations in which one group feels threatened by another group, and thus preemptive conflict ensues. These situations arise if one group has a geographic advantage over the other. This can happen, for example, if a group A has "islands" of its kin scattered across the territory of another group, B. These islands consist of small populations of A with a low infrastructure and would thus be easy to conquer by B. As a result of this perceived threat by B, A is likely to take military action to defend its ethnic islands. In summary, groups that settle in single, contiguous clusters are less likely to engage in conflict, since the mutual dependency created by island population patterns does not exist. Fearon (1998) makes a similar claim. In his formulation of ethnic conflict as a commitment problem which I introduced above, he predicts that secessionist claims by a minority will not result in conflict if the minority occupies a contiguous piece of land such that territorial separation is feasible.

In general, the impact of intermingled settlement patterns on conflict can go both ways. As Posen and Fearon suggest, clear territorial boundaries with a low degree of ethnic mixing are conducive to peace. This is in line with Kaufmann's suggestion of territorial separation

as a solution to end ethnic violence (Kaufmann 1996), who claims that in a state of hardened ethnic boundaries and continuing hostilities between groups, the only way to end violence is to spatially separate groups. Quite the contrary, intermingled settlement patterns have also been mentioned as supportive for peace. Here, the reasoning is that if groups live together at the local level, this leads to frequent social interaction between groups and thus reduces tensions.

However, even though existing explanations have a lot to say about how interconnected group geographies affect the risk of conflict, none of these approaches can tell us *where* group conflict is likely to occur. Under a security dilemma perspective, are the small territorial islands more susceptible to violence? The location of inter-ethnic violence has not received much attention in the literature. Below, I introduce the foundations for a theory on the location of ethnic violence, to be explained in detail in Chapter 6.

Groups, Territory and Boundaries

Numerous authors have highlighted the importance of territory for groups. In Smith's definition of an ethnic group, the attachment to a given piece of land is a crucial part of the group's identity (Smith 1986). The feeling of belonging that unites members of an ethnic group not only rests on a common origin, shared language and/or religion, but also has a spatial component: An ethnic group defines itself as the people that belongs to a particular region. This assumption is central to Toft's "Theory of Indivisible Territory", in which she assumes that the perception of a territory as "homeland" leads the group to demand complete control over it (Toft 2002, 2003).[4]

Where does the group territory begin, and where does it end? Existing accounts of ethnic territory, such as Posen or Toft, have a tendency to take the extent of a group territory as given. In some cases, this is a valid assumption since the spatial boundaries of a group territory often correspond to existing internal administrative boundaries. For example, the territory of Wales – even though it is now part of the United Kingdom – can still be recognized since it constitutes one of the first-level administrative units within the UK. However, not all ethnic territories are well-defined. Consider the example of the former Yugoslav Republic of Bosnia, where Croats, Serbs and Bosniaks lived in highly intermingled settlement patterns.

Under normal circumstances, the precise boundaries of group territories are of little importance. However, during the break-up of multi-ethnic states, group boundaries can become subject to intense negotiation and even fighting between ethnic groups. This condition has been described by Posen (1993*a*) as the ethnic security dilemma under a temporary state of anarchy between groups. During the break-up, groups face the prospect of creating new state entities

[4]However, see Goddard (2006) who argues that territorial indivisibility is not a constant, but depends on how actors legitimize their claims.

from what used to be a common state, and the natural question is what these new units will be. I assume that during this process, conflicting claims of particular groups to certain locations drive the application of violence: If there is only one group demanding control over a particular piece of land, it can be added to the group's territory without resorting to violence. On the other hand, if two or more groups claim ownership of a location, under the state of a temporal anarchy there is no central authority that can resolve this conflict peacefully. Instead, groups will try to assume control by military means.

Which locations are of particular importance to groups? There are two considerations we must take into account. First, ideally the entire group population should be included in the state to be created. In other words, the internal boundaries along which the former union state is split should be drawn such that the new states include all members of the respective groups. Second, each state must have a feasible territorial configuration. As a requirement, the state's territory should be contiguous. It is obvious that a state with territorially isolated enclaves will face enormous difficulties in defending them (Posen 1993*a*; Kaufmann 1996). Also, a contiguous territory assures economic benefits because of uninterrupted communication and transportation routes (Melander 2007). I assume that these two considerations together create strategic incentives for groups to fight for locations that have a high proportion of their kin, but at the same time can easily be added to the larger new state.

Summary: Territorial Configurations and Strategic Ambitions

Settlement patterns of ethnic groups create interaction opportunities between groups, which determine the location of ethnic violence. I have argued that during the break-up of a multi-ethnic state, groups fight with the prospect of creating new state entities from their group territories. If group territories are clearly demarcated, secession will be feasible. However, if groups live in intermingled settlement patterns, the geographic shape of the state entities is subject to violent negotiation by groups. In this section, I have laid out the foundations for a theory that explains where this violence is applied. Essentially, I assume that groups follow two considerations: future state entities should ideally include all of a group's kin, but should also form a contiguous territory in order to make the state feasible to maintain and defend. These strategic considerations lead groups to regard certain locations as particularly important. Conflict over a location arises if it is considered important for more than one group. Chapter 6 provides a disaggregated empirical analysis of these propositions.

2.5 Chapter Summary

In this chapter, I have outlined the theoretical foundations for the work presented in this dissertation. I started with a definition of ethnic civil war as an internal conflict between a government and an ethnic group. I discussed the various difficulties associated with "ethnic groups" as an analytical category, distinguishing between essentialist conceptions of ethnicity that see ethnic boundaries as constant and exogenously given, and constructivist approaches, which emphasize the social construction of group boundaries and their fluidity. The theoretical approaches in the literature that explain ethnic conflict can be categorized into two types: a symmetric perspective of conflict between groups of equal status, not necessarily in the context of a state; and an asymmetric perspective of a minority group discriminated against by the state.

Geographic settlement patterns influence ethnic conflict at two levels. At the micro-level, I take a monadic perspective and assume that the spatial distribution of a group provides interaction opportunities for group members. Social interaction between group members is conducive to conflict since it fosters the establishment of collective grievances, facilitates the flow of information, and increases social pressure on individuals to participate. If group geography is such that it facilitates interaction, a group should be easier to mobilize for collective action and thus carries a higher risk of conflict. Settlement patterns also influence conflict at the meso-level by providing interaction opportunities between groups. During the break-up of multi-ethnic states, groups try to consolidate their territory by means of violence. Violence at a particular location will result if two groups consider that location as important enough to be included in the new state.

Chapter 3

Spatial Data on Ethnic Groups

ANALYZING group settlement patterns introduces a spatial dimension to the study of ethnic conflict that sets it apart from much of the previous work on this phenomenon. Not surprisingly, the processing of spatial data as required for this dissertation has traditionally been a domain of the geographic sciences. It is therefore useful to look beyond disciplinary boundaries and borrow some of the concepts and techniques geographers have developed. This chapter gives a short introduction to the basic concepts of spatial data collection and processing using Geographic Information Systems (GIS) procedures. I then present the two primary data sources this dissertation builds on: First, the *Geo-referencing of Ethnic Groups* (GREG) dataset that covers ethnic groups around the globe, but at a rather coarse level of aggregation, and second, high-resolution settlement pattern data obtained from census data, available for a limited region only.

3.1 Existing Data on Group Settlement Patterns

This chapter introduces two data sources for ethnic group settlement patterns, one of which was created from scratch. Why bother developing a new dataset? The answer lies in the simplified data representation employed by other datasets that makes them unsuitable for studying mechanisms of conflict. In the following, I will formulate three requirements the data for this dissertation must satisfy in order to answer the research questions. I will show that existing sources fail to meet these requirements, which is why I have resorted to other sources of information.

The first core question of this dissertation studies how the settlement pattern of a particular group affects its likelihood of being involved in conflict. In the previous chapter, I argued that settlement patterns determine the interaction opportunities of group members. Successful

group mobilization for conflict is more likely if the settlement pattern favors frequent and close interaction between individuals. In particular, I argued that social interaction is more likely if individuals are located close to each other. The proposed relationships will be spelled out in more detail in the following two chapters. However, already at this point it is obvious that for an empirical test of the proximity–interaction relationship, we require detailed information on the location of group members relative to each other.

The second core question is how settlement patterns of groups relative to each other affect the location of violence. Here, the focus is much more narrow than in the previous question: instead of studying conflict involvement in general, I aim to explain the occurrence of violence at the level of events. In particular, as I will outline in Chapter 6, I examine how the degree of local ethnic mixing and the location of a particular unit relative to larger group territories affect the propensity for violence. This entails a second data requirement: The resolution of the settlement pattern data needs to be sufficiently fine-grained so that we can observe local variation in the ethnic configuration that may potentially be related to the application of violence.

Last, the third core question takes a dynamic perspective and traces the relationship between conflict and settlement patterns. Settlement patterns not only affect whether and where violence is applied, but also vice versa: The spatial arrangement of ethnic group members is likely to change as a result of conflict. Clearly, for a systematic study of this endogenous process, we require spatial data on groups over time. In summary, for the analysis presented in this dissertation, the data must show where group members are located relative to each other, what the degree of local ethnic mixing is, and how ethnic settlement patterns change over time. Can existing data sources satisfy these requirements?

One of the few (if not the only) systematic data collections that includes geographic variables of groups is the *Minorities at Risk* dataset (Minorities at Risk Project 2005; Gurr 1993). MAR is a worldwide list of groups facing discrimination and includes several variables describing the groups' settlement patterns.[1] Most importantly, however, MAR codes the degree of concentration of the group settlement pattern, distinguishing between four major types of groups: those that are "widely dispersed", "primarily urban", "majority in one region", and "concentrated in one region". If applicable, the dataset also contains the group's population shares in its regions or cities.

The rough coding in MAR makes it unsuitable for the project at hand. First, as I have outlined above, in order to test spatial mechanisms of group mobilization as proposed in the previous chapter (and later to be expanded in Chapters 4 and 5), detailed information on the

[1] Here, I refer to the "old" MAR group concentration variables used e.g. in Toft's (2003) group settlement pattern analysis. Even though the "new" concentration indicators are an improvement over the old ones, in principle they suffer from the same shortcomings I describe here.

location of group members is required. MAR captures the settlement pattern with a few numeric indicators that do not allow inferences about the location of group members relative to each other. By the same token, MAR fails to meet the second data requirement – the availability of information on degrees of spatial mixing between groups. By construction, MAR codes at the level of groups, so if the group shares a certain region with another group, we do not know which groups it is. Third, whereas MAR codes group attributes over time in principle, the rough coding of settlement patterns makes it impossible to trace geographic changes during conflict. These changes would only become visible in the MAR data if they were so fundamental that there was a shift from one MAR concentration category to another (e.g. from "regional majority" to "concentrated"). However, much of the dynamics during conflict occurs at a finer level and thus fails to show up in MAR.

More fundamentally, the MAR dataset suffers from a problem of selection bias (King, Keohane and Verba 1994). The inclusion criterion for groups to MAR is that groups that suffer from systematic discrimination vis-à-vis other groups, or collectively mobilize for political goals. Obviously, this is problematic if we are to study the effect of certain group characteristics (such as settlement patterns) on conflict, since MAR explicitly selects groups with a potential risk of conflict. What is required is a more inclusive worldwide list of groups that allows us to compare groups with respect to their settlement patterns, without limiting the scope to groups that experience discrimination. In summary, MAR would not be a suitable data source for this project, which is why I resort to two other datasets, described below. Before turning to these datasets, I briefly introduce a way of representing group geography which improves upon the limited MAR procedure of coding settlement patterns with numerical indicators described above.

When creating a numeric indicator, we make a small part of reality operational for quantitative analysis. In the example of group concentration mentioned above, what we are interested in is whether a group is concentrated. However, concentration represents only a small aspect of the spatial group distribution. Alternatively, we could measure whether the group territory borders another group, or the degree of ethnic mixing in the group's home region. All these variables, however, have in common that they represent particular aspects of the same underlying dimension, the distribution of groups in a country.

The "non-spatial" coding scheme of representing concentration by a 0/1 indicator seems at first intuitively plausible. However, at the same time it is very limited in scope. What if we need an indicator that is not contained in the dataset? For example, we might be interested in the distance between the population clusters of a group (as I demonstrate in Chapter 4), but there is no such variable in the MAR dataset. For that reason, it is recommendable to resort to a different approach to data collection than the coding at the indicator level described above. Rather than

coding *indicators* of the settlement pattern, in a "spatial" coding scheme we represent the spatial distribution of groups directly in an (electronic) ethnic map. Even though data collection and representation will look slightly more complex, the advantage is that we can now compute many different indicators directly on these maps. Beyond the benefits of a flexible use of the data, this more advanced coding procedure also increases transparency as it makes explicit the geographic information contained in the indicators. For example, suppose we wonder why, according to the MAR dataset, a particular group X in country Y has a concentration value of 1. In order to find out, we would have to search for a map that shows the settlements of X in Y. If we code the ethnic map directly, this question can easily be answered since we have the geographic information available at hand.

In summary, despite its higher complexity, the spatial coding scheme has many advantages over simple indicators. Especially for this dissertation, which requires information about different aspects of a group settlement pattern, it is the only way forward. This of course has major implications since a geographic dataset of groups does not exist and will have to be created from scratch. The remainder of this chapter describes the creation of this dataset. However, before turning to the dataset itself, I briefly describe the GIS technology used for its creation.

3.2 Representing Group Geography in GIS

In recent years, Geographic Information Systems (GIS) have become increasingly popular beyond the boundaries of the geographic disciplines and have also been applied to social science problems. GIS allow for the collection and processing of spatially-anchored data using a computer and are therefore well-suited for the study of group settlement patterns. Although mostly referred to as such, the term "GIS" does not just refer to the software used to process spatial data. Rather, GIS is used more broadly and encompasses the data formats, algorithms and map creation techniques required for the electronic processing of spatially referenced data. This section gives a brief introduction to the representation of spatial data in a GIS.

3.2.1 GIS Data Models: Vector vs. Raster

GIS data come in two fundamentally different formats – vector and raster data. As the following paragraphs explain, these two formats have different strengths and weaknesses, and therefore complement each other well. For many GIS analyses, it is necessary to combine data from both formats, so almost all of the standard GIS tools can process both of them.

GIS *vector* data are used for discrete spatial entities, e.g.cities, rivers, or countries. The entities are represented in the GIS as points, lines, or polygons – whatever fits the respective

geographic entity best. For example, villages are typically stored as single points, so-called *point* features. Rivers can best be represented as *line features*, where each line consists of a series of concatenated points. Lastly, for any kind of area as for example a lake or a country, a *polygon* feature is used. A polygon is a series of concatenated points in which the first and the last point are the same. Figure 3.1 shows a graphical illustration of the three vector feature types.

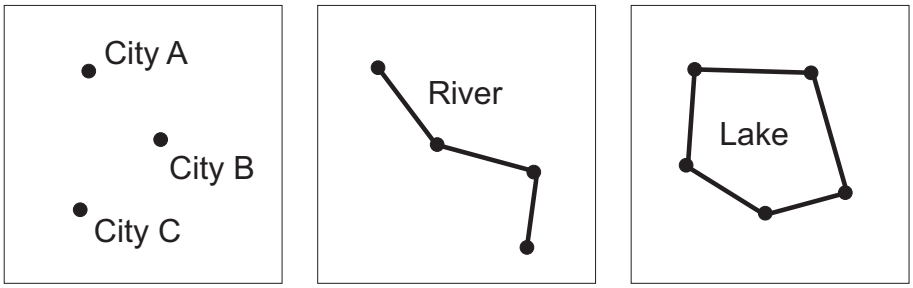


Figure 3.1: GIS vector data formats: Point features are typically used for cities and villages (left), rivers and roads are represented by line features (center), and areas such as lakes are stored as polygon features (right).

A GIS vector representation is the format of choice when dealing with clearly delimited, separate geographic entities. However, what if we need to store information about a variable that is spatially continuous, i.e. where a value exists for each and every point in space? For example, it would be difficult to store territorial elevation data in a vector format, since there is one elevation value for every location on earth. In a vector format, we would have to create a large number of point features, each of which would have its own elevation value. However, the second GIS data representation I introduce is better suited to dealing with this type of data. In a GIS *raster* representation, the geographic space is divided into equal-sized, quadratic cells. For each of these cells, the dataset contains a numeric value, as is the case for elevation, illustrated in Figure 3.2.

All GIS data formats – vector and raster – require the specification of the geographic space they operate on. In short, we need to define how geographic locations from the earth's three-dimensional surface correspond to locations in the two-dimensional map. In earlier times, these geographic reference systems were normally defined for individual countries. However, in the last decades a worldwide reference system was developed that makes it possible to use the same system across national boundaries. This system – the World Geodetic System 1984 (WGS84) – now constitutes the de-facto standard in GIS applications. Unless otherwise noted, the data I use are WGS84-referenced, with geographic (latitude/longitude) coordinates.

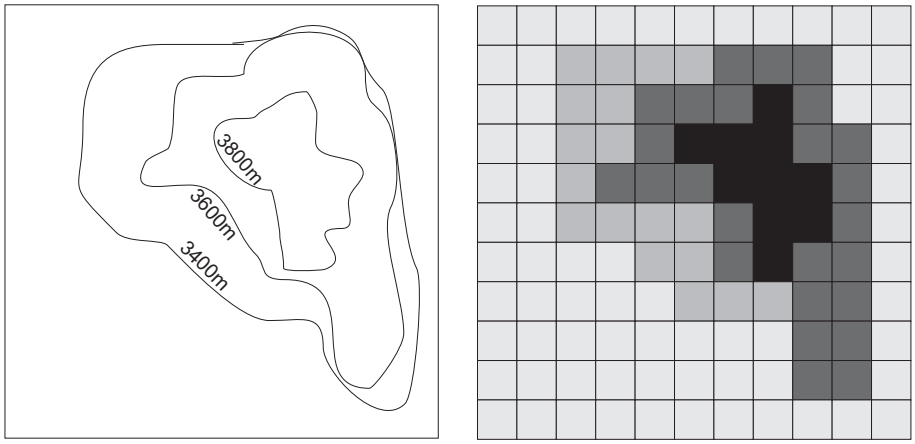


Figure 3.2: GIS raster data format: The geographic space is divided into equal-sized cells, each with its own numeric value. The example shows how territorial elevation (shown by the contours, left) can be represented as a raster (right, darker colors correspond to higher elevation values).

3.2.2 Spatial and Non-spatial Information in a GIS

The datasets on ethnic settlement patterns presented below rely exclusively on the GIS vector format. Unlike GIS raster data, a vector dataset can contain both spatial information (the features: cities, rivers, or lakes) and non-spatial information. The non-spatial information is contained in a data table with a set of variables, similar to other tabular data formats used in statistics. This table is called the *attribute table*. The lines in the attribute table are connected to individual geographic features, i.e.spatial entities represented in the dataset such as cities, rivers or lakes, so that each line in the attribute table corresponds to exactly one feature. These features are stored with their precise coordinates in a given projection and coordinate system. For example, a vector dataset could consist of a set of points representing cities, with the non-spatial information about these cities (city name, population etc.) stored in the corresponding records in the attribute table. Figure 3.3 shows an illustration.

3.3 The GREG Dataset[2]

The remaining sections of this chapter describe two GIS datasets of ethnic settlement patterns used for the analysis presented in this dissertation. This section introduces the first of them, the *Geo-referencing of Ethnic Groups* (GREG) dataset. As I outlined above, GREG is a spatial dataset which uses a GIS data format to allow for sufficient flexibility in computing spatial

[2]This section is based on Weidmann, Nils B., Jan Ketil Rød, and Lars-Erik Cederman. 2008. "Representing Ethnic Groups in Space: A New Dataset." Under review.

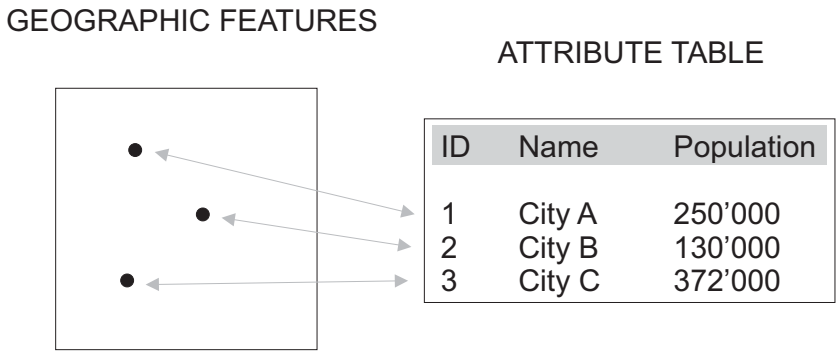


Figure 3.3: Storage of non-spatial information in a GIS vector dataset: Each feature is associated with a line in the attribute table that contains supplementary information about the feature.

indicators. I start with a discussion of the data sources considered for GREG and then turn to a description of the coding procedure and the dataset format. I close this section with a description of the shortcomings of the dataset.

3.3.1 Data Source: The Atlas Narodov Mira

In the literature, ethnicity has typically entered analysis of conflict processes either as qualitative, historical entities or as quantitative indices, such as the ethno-nationalist fractionalization index (ELF, Fearon and Laitin 2003; Posner 2004). In neither case, however, has space played a prominent role in the development of causal arguments. Historical and other qualitative accounts of ethnic conflict occasionally provide maps that show the spatial distribution of groups, but such information is hardly ever supplied for a larger sample of states (Horowitz 1985; Herbst 2000; Kocher and Kalyvas 2003). While the quantitative literature does offer some references to group geography such as settlement concentration, this information is usually narrowed down to a few variables (e.g. the Minorities at Risk Project, mentioned above). To my knowledge, however, there are no data sources that systematically pinpoint the location of ethnic groups in a large number of comparable cases. Thus, the exact location of the key actors in ethnic conflicts is still obscure. This raises the question where such information could be drawn from. Several candidates exist.

Linguists have developed detailed maps of language diffusion, for example *Ethnologue* (Gordon 2005). However, this data resource is problematic because it is narrowly focused on linguistic traits. Its linguistic charts are typically either too detailed to serve as a guide to ethnic group delimitation or too sketchy as they often represent a linguistic groups with a point, thus making delimitation or inference on spatial dissemination fuzzy. Moreover, *Ethnologue* does not include spatial information for all countries, making it unsuitable for the development of a

global dataset.

Another possibility is to infer the location of ethnic groups from census or survey data. Yet such an approach is only viable where such data contains references to ethnicity, which is often not the case. Furthermore, this approach hinges on the presence of a reasonably fine-grained provincial structure. Where federal subunits are large, the needed degree of spatial disaggregation may never be attained. A few cases exist for which this approach is possible. For example, the 1991 census for Bosnia provides detailed information about the ethnic composition of municipalities (Petrovic 1992, see below). In most cases, however, providing spatially referenced census data for a larger set of cases is not possible.

For the aforementioned reasons, I have chosen to rely on data and maps from the well-known *Atlas Narodov Mira* (ANM, Bruk and Apenchenko 1964), which stems from a major project that charted ethnic groups, undertaken by Soviet ethnographers in the 1960s. Their efforts yielded the extensive but still un-translated ANM that covers the entire globe. The ANM has several strengths: It is complete and carefully researched, it relies on a uniform group list that is valid across state borders, and it provides high-quality maps. At the time of its publishing, the ANM received excellent reviews (Hewes 1966). Among its advantages, Harris (1965) explicitly mentions the inclusion of minority groups in states dominated by other groups, which makes it particularly suitable for the study of ethnic conflict. Until now, the ANM has been widely used in contemporary research, mainly as a basis for calculating the ELF index (Taylor and Hudson 1972).

3.3.2 Creating a GIS Dataset from the ANM

In this section, I describe the efforts to create a GIS dataset on ethnic groups on the basis of the ANM. I start with a detailed introduction of the ANM data and then show how this information was converted into GIS format. The ANM consists of 57 ethnographic maps, covering all regions of the world at various scales. Each map shows the geographic distribution of the relevant groups, indicated by colored areas. In addition to the color coding, the areas are marked with numbers which refer to the respective group's number in the legend. Most areas are coded as pertaining to one group only, but in some cases there are up to three groups which share a certain territory (although the latter case is quite rare, see below). This is indicated on the map by a striped pattern. Figure 3.4 shows a part of the map covering former Yugoslavia. Note that for each area there are one or two numbers indicating the respective group.

The ANM also provides information about groups without a clear territorial base. The presence of these groups is indicated by symbols rather than areas. To give an example in the above map, minor occurrences of group 4 (Croats, square symbol) can also be found in northeast

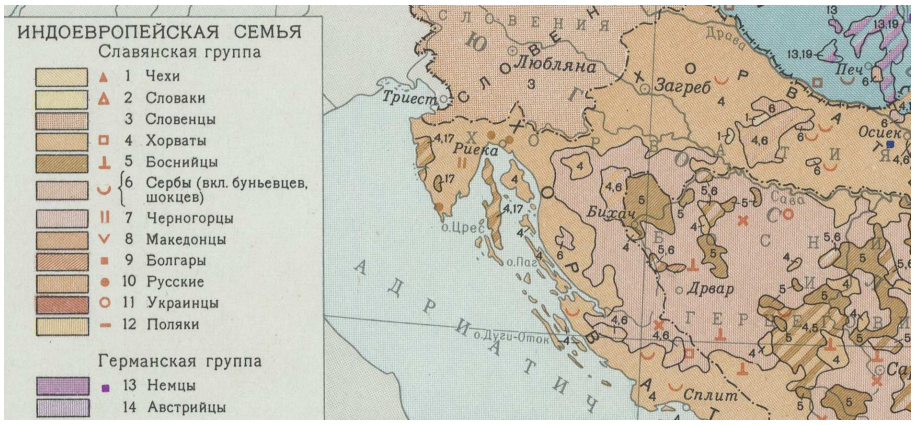


Figure 3.4: Map of the former Yugoslavia from the Atlas Narodov Mira.

Slovenia. Sparsely populated areas can be distinguished from others by their grey raster and the missing group color. However, for these regions group presence is still indicated by symbols and numbers as explained above. Unpopulated regions are left white.

The source of the information contained in the ANM remains somewhat unclear. A short text at the beginning of the volume list three different types of sources: (1) ethnographic and geographic maps assembled by the Ethnography Institute of the USSR Academy of Sciences, (2) population census data, and (3) ethnographic publications of government agencies. Still, it remains unclear what kind of information was used for which maps. Since the type of source determines the resolution at which we obtain ethnographic data, a comparison of data from different sources might be problematic. This is an issue to which we will return later. Apart from its map collection, the ANM features a statistical appendix which complements the geographic information. The appendix contains two major lists. The first one presents the full set of groups mentioned in the ANM along with their relative population sizes for each country, and the second contains all countries together with their groups. It is the latter list that has served as a basis for the computation of ELF scores in the literature (Taylor and Hudson 1972).

The aim of the GREG project is to make the ethnic maps usable for the spatial analysis on ethnicity. To this end, the ethnic group regions marked in the ANM were converted into a GIS dataset. Groups without a clearly defined territory (those marked with symbols on the map) were not included. For the dataset creation, three steps were carried out. First, the maps from the ANM were scanned to obtain an image file for each map. Second, these images were then spatially referenced using a geographic information system (GIS). More precisely, using easily identifiable geographic points such as cities, river bends or border junctions, each image was aligned to the GIS's coordinate system such that all points on the map matched their true geographic coordinates. Once a satisfactory match had been obtained, the group areas were

redrawn in the GIS, thus yielding one or more vector polygons for a group.

Since all the maps in the ANM are annotated in Russian, in a third step the group names in the ANM were assigned by a native Russian speaker. Translations of the Russian group names into English are already given in the ANM's appendix, so this step consisted of matching group polygons with the appropriate translation in the English group list. The final result of this merging step and thus of the GREG project is a master list of ethnic groups, each with a unique numeric identifier, and a set of polygons in ESRI's shapefile format (ESRI 1998), with three attributes: G1ID, G2ID and G3ID. These attributes store the identifier of the ethnic group(s) represented by the given polygon. For polygons with only one ethnic group, only G1ID is used, and the value of G2ID and G3ID is set to 0 (no group). Correspondingly, for polygons with only two groups the G3ID attribute takes the value 0. Figure 3.5 shows the geo-coded groups from Figure 3.4.

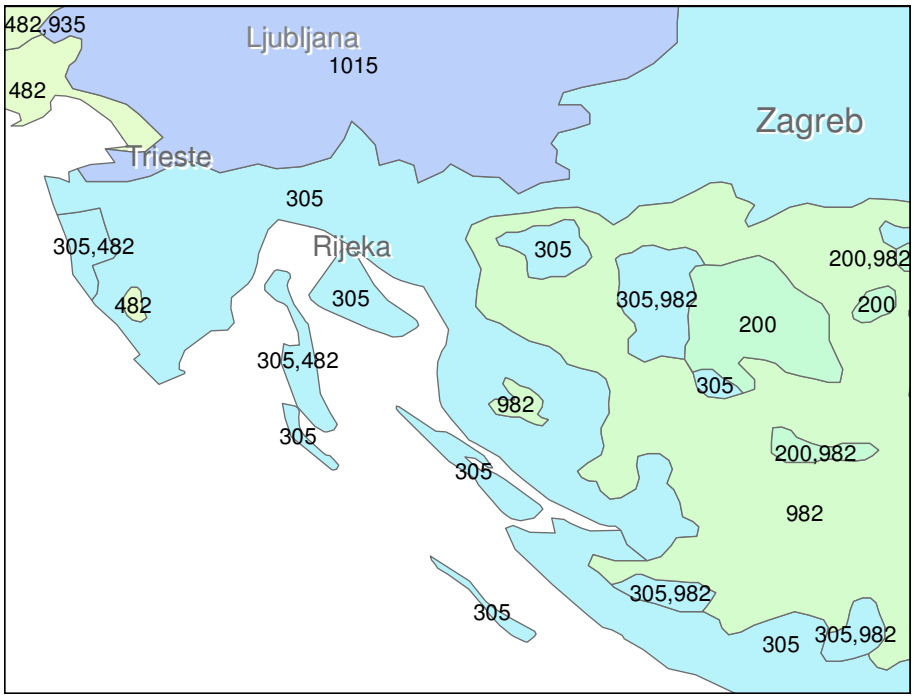


Figure 3.5: The geo-coded group polygons from Figure 3.4. The numbers are unique identifiers for the groups, with some polygons belonging to more than one group. Colors of polygons are assigned by the first group (G1ID).

The full GREG dataset has global coverage and consists of 929 groups represented by 8,969 georeferenced polygons. In the ANM, there are 1,248 groups in total but as 319 of these do not have any clearly defined territory, they are excluded from the GREG dataset. Of the 8,969 polygons, the vast majority (7,383) contains one group. 1,552 polygons contain two groups and for only 34 polygons, there are three groups given in the dataset. The size of the polygons varies

considerably: The smallest polygon occupies an area of 0.59 square kilometers, and the largest polygon extends over 6,954,564 square kilometers.

3.3.3 Potential Problems of the ANM Data

Outdated

The data source for GREG is the Atlas Narodov Mira, whose data was collected in the early 1960s. For this reason, an obvious problem with the GREG data is that it might no longer capture the present ethnic configuration. Still, ethnic settlement patterns exhibit a lot of inertia, so that it is plausible to also use the GREG data as the basis for measuring ethnic geography in recent times. Especially during conflicts, however, ethnic configurations might change significantly, for example by conflict-induced population movements, the systematic expulsion of people, or even ethnic cleansing. In order to capture these changes more accurately, one would have to introduce a time dimension to GREG. This could be done by providing geographic data for multiple periods during which the settlement pattern of a group remained roughly constant. The Bosnia dataset introduced below follows this approach, but for a more limited geographic region.

Group Categories Problematic

What is an ethnic group? Ethnic groups can be distinguished along linguistic or religious lines, with or without a fixed territory, or according to their political relevance. Due to the coding conventions of its data source, the GREG dataset distinguishes groups primarily according to their language. This is an important caveat to keep in mind when using GREG. A short comparison of GREG group categories with those employed in other datasets illustrates this point. The first dataset I use for comparison is the list of relevant ethnic groups compiled by Fearon (2003). Second, I compare the GREG list to one of the most widely used datasets on ethnic groups, the MAR dataset introduced above (Minorities at Risk Project 2005; Gurr 1993).

The comparison of group lists is a difficult task because in many cases, group names are not consistent and thus need to be matched manually. For that reason, I limit my efforts to four countries: Belgium, Iraq, Georgia and Cambodia. Table 2 summarizes the comparison. Shaded lines list the matching groups in the three datasets. Empty cells indicate that no match was found in the respective dataset.

Belgium For Belgium, GREG and the Fearon dataset list the major groups of the country (Flemings and Walloons). However, there is disagreement on what minority groups should

Country	GREG	Fearon (2003)	MAR
Belgium	Flemings	Flemings	
	Germans		
	Walloons	Walloons	(none)
		Italians	
		Maroccans	
Iraq	Assyrians		
	Circassians		
		Shi'ites	Shi'ites
	Iraq Arabs		
		Sunni-Arabs	Sunnis
	Kurds	Kurds	Kurds
	Lur		
	Persians		
	Turkmens	Turkomans	
Georgia	Abkhaz	Abkhazians	Abkhazians
	Armenians	Armenians	
	Azerbaijanians	Azeris	
	Bats		
	Estonians		
	Georgians	Georgians	
	Greeks		
	Ingushes		
	Moldavians		
	Ossetes	Ossetians (South)	Ossetians (South)
	Russians	Russians	Russians
		Adzhars	Adzhars
Cambodia	Boloven		
	Cham	Chams	Chams
	Jarai		
	Khmers	Khmers	
	Kui		
	Lao		
	Ma		
	Malays of Malaya		
	Muong and Brao		
	Siamese		
	Stieng		
	Vietnamese	Vietnamese	Vietnamese
		Chinese	

Table 3.1: Comparing the GREG, Fearon (2003) and MAR group lists for four countries.

be included. GREG includes the small German-speaking population in eastern Belgium. It is unclear why the Fearon dataset excludes this group, given that it fits most of his criteria for an ethnic group (Fearon 2003, p. 201). On the other hand, Fearon includes Italians and Moroccans who are most likely foreign workers in Belgium with limited political relevance. These groups, however, do not have a territorial base in Belgium (and were probably not present at the time the ANM was created), so they are not included in GREG. MAR does not list any groups for Belgium, since none of the groups are seen as experiencing discrimination.

Iraq The ANM's focus on linguistic boundaries poses problem for cases where alternative group distinctions have become relevant. An example is Iraq, where divisions between different Muslim denominations – Shi'ites and Sunnis – have dominated the country for a long time. This distinction is made both in the Fearon dataset and in MAR. However, the GREG dataset summarizes these groups under a single identity, "Iraq Arabs", since both speak Arabic. Sunnis and Shi'ites differ significantly in terms of their spatial distribution, with the Shi'ites dominating the south of the country, and the Sunnis mostly present in the north of Baghdad and in the west. It would therefore be possible to pin down the relevant groups geographically, but one would have to rely on alternative data sources. Again, MAR's selection of minorities facing discrimination means that governing ethnic groups are excluded from the dataset. GREG includes smaller groups that speak a different language, for example the Assyrians, Circassians, and Persians. The Kurds are included in all three datasets. The Turkmens of Iraq, however, fail to be listed in the MAR dataset.

Georgia As in the case of Belgium, GREG and the Fearon dataset agree with respect to the most important groups in Georgia (Georgians, Azeris, Armenians, Russians, Abkhazians, and Ossetians). However, the Fearon group list distinguishes the Adzhars (or Ajars) from the rest of the Georgians because of their different religion (Islam), whereas the ANM does not make this distinction. This might be due to the radical suppression of Islam during the Soviet era (Hughes and Sasse 2002), where the ANM criteria for inclusion might have been guided by political aspirations. The ANM, however, lists other minority groups without political relevance. Four groups are regarded as being discriminated and are therefore included in the MAR dataset, including the Adzhars whom the ANM fails to list and the secessionist Abkhaz and South Ossetians. By definition of MAR's coding rules, the titular majority – the Georgians – fails to show up in the MAR group list.

Cambodia Among the four examples described here, Cambodia shows the highest level of detail in the ANM. The GREG dataset contains detailed information about the highland tribes

(Jarai, Kui, Muong, Stieng) which is present in none of the other datasets. GREG and the Fearon dataset agree on the major groups (Khmer, Vietnamese, and Cham), but GREG does not list the Chinese, which constituted a significant minority already in the 1960s (Ross 1987). This is due to the fact that the Chinese were engaged in commerce all across Cambodia, so they do not have a delimited territorial basis in the country.

In sum, our comparison highlights major differences between datasets on ethnic groups. MAR by definition shows only parts of the ethnic landscape by focusing on discriminated groups only. In the case of Iraq, the ANM's focus on linguistic boundaries causes GREG to omit the Sunni–Shi'i division, one of the most important cultural cleavages in the country. Since the coding criteria for the ANM are not spelled out in detail, in some cases it is difficult to reproduce the coding decisions made. Most of the distinctions are clearly based on linguistic differences. Additionally, national boundaries were introduced. For example, the ANM distinguishes between Germans and Austrians, even though they belong to the same language family. Similarly, the ANM separates "English Irish" from "Irishmen". These coding decisions are problematic in some cases and require the additional inspection of the group list. Furthermore, it is important that as a spatial dataset, GREG only includes territorial groups, i.e. those that have one or more settlement regions in a country. Migrant groups or foreign workers residing only in urban areas are not represented in GREG.

Different Resolutions of Spatial Data

As mentioned above, different data sources were used to compile the ethnic maps in the ANM. Most likely, detailed census data were used for areas under Soviet influence, whereas data for other countries were provided at much coarser resolutions. The example of China and Russia demonstrates this. In both countries, the distribution of the total area occupied by a group (i.e. the sum of all the group polygons it occupies) is roughly comparable. The log-transformed average group area (measured in square kilometers) is about 9.23 in China, and 9.26 in Russia. In a two-sample t-test, the two distributions are not significantly different (t = 0.06, df = 113.8, p-value = 0.95). However, the areas of the individual group polygons differ: Figure 3.6 shows histograms of the area of the GREG polygons in the two countries. In Russia, we find a large proportion (about 40%) of smaller polygons with an area smaller than 10,000 square kilometers, whereas in China these polygons only amount to 10% of the total number of polygons. A t-test confirms this (t = -15.49, df = 1223.49, p-value = 0.00). In sum, these results show that the geographic information on groups is provided at different resolutions in China and Russia. Depending on the respective application, this will limit the comparability across countries. In

the remainder of this dissertation, I address this problem by weighting polygons according to their area or population (see Chapter 4).

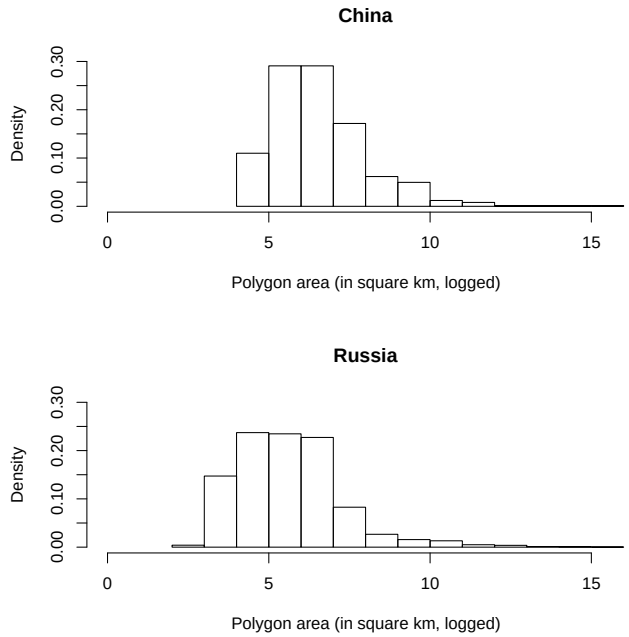


Figure 3.6: Distribution of polygon areas for GREG polygons in China and Russia.

3.4 Bosnia: Spatial Census Data

The GREG dataset introduced above has several advantages that make it particularly useful for cross-national comparisons: Most importantly, it includes data on almost every country on the globe, and it employs the same group categorization across countries. However, the large geographic scope of the data comes at the expense of detail. GREG divides countries into the major settlement regions of groups, but is unable to give the local degree of mixing for a particular location. Furthermore, as mentioned above GREG does not trace *changes* in settlement patterns, as they typically occur during conflict. For these reasons, I created a second dataset for the purpose of this dissertation. By relying on census data as the source of information about geographic settlement patterns, I achieve the high level of detail on the population distribution that GREG lacks. Furthermore, this approach allows me to obtain time-series data on settlement patterns and thus study the endogeneity of settlement patterns in conflict. However, the downside is that the required census information is only available for a limited region. Bosnia satisfies most of the data requirements for the corresponding chapters and was thus chosen as

the area of study.

3.4.1 Census Data for Bosnia

Data on Bosnia's *pre-war* ethnic distribution is available from the Republic of Yugoslavia's last census in 1991 (Petrovic 1992). For each of the 109 municipalities (opštinas) in Bosnia, the census lists the number of people from the three major groups: the Bosniaks or Bosnian Muslims, the Croats and the Serbs. These 109 municipalities from the 1991 census constitute the unit of observation in the Bosnia dataset. They allow for the analysis of conflict dynamics at a high level of resolution: Whereas most quantitative works on internal conflict employ a state-level measurement, using the Bosnia data we can narrow the focus down to municipalities of (on average) 20x20 kilometers.

According to the 1991 census, the majority of the Bosnian municipalities are Bosniak-dominated: Of the 109 units, 40 have a Bosniak majority, 33 a Serb- and 15 a Croat majority. However, the census data shows that there was a significant degree of ethnic mixing in the Bosnia of 1991. One way to measure ethnic mixing is by means of the Ethno-linguistic fractionalization index (ELF Taylor and Hudson 1972; Alesina et al. 2003). The ELF index f is computed for each municipality on the population shares s_i of the three major groups i, according to the following formula:

$$f = 1 - \sum_i s_i^2$$

The range of f on a sample with a maximum number of three groups per municipality is from 0 to 0.66, where 0 corresponds to a unit where there is one group with population share 1.0 (high homogeneity), and 0.66 for a unit with three equally-sized groups (high heterogeneity). The plot in Figure 3.7 shows that the majority of the municipalities in Bosnia are close to the maximum ELF value.

In order to analyze the *changes* in settlement patterns over time, I also obtained post-war estimates of groups at the municipality level. These numbers are taken from Caspersen (2004) where election results from 2000 were used to approximate the number of people of a particular ethnicity in a unit. Since many of the old municipalities were split after the war, I aggregate the post-war population data to the pre-war municipalities. This is done using pre-war and post-war maps to identify which of the new units correspond to the 1991 census units.

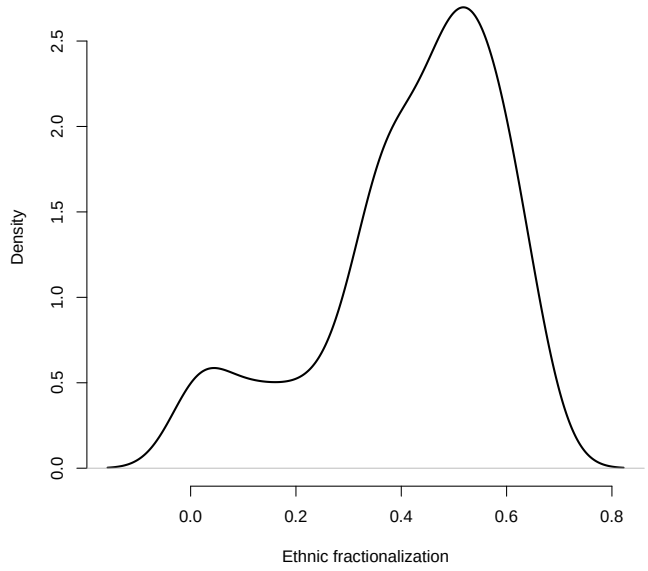


Figure 3.7: Distribution of ELF scores for the Bosnian municipalities, smoothed using a kernel density estimator. The highest possible ELF score is 0.66 (because of the kernel density estimation, there appears to be a non-zero density for values above this limit). The plot shows the high degree of ethnic mixing in Bosnia in 1991.

3.4.2 The Spatial Extent of the Bosnian Municipalities

The population figures at the municipality level needed to be combined with geographic information about the relative location of the units. For that reason, I created a GIS vector dataset of the pre-war municipalities that contains the detailed boundaries.[3] The area of the municipalities ranges from about 32 square kilometers for the smallest municipality (the *Centar* district of Sarajevo) to 1,300 square kilometers for the largest municipality (Srbinje in the south-east of the country), with a mean of 466 square kilometers.[4]

Once the geographic location of the municipalities is determined, it is possible to map the ethnic distribution at the municipality level described above. Figure 3.8 shows the settlement patterns of the three major ethnic groups in Bosnia: the Bosniaks, Serbs and Croats. The figure shows that there are different primary settlement regions of the groups in Bosnia, with the Bosniaks being located primarily in the North and South, the Serbs in the East and the Croats in the West.

[3]Thanks to Jan Ketil Rød for this help in creating the municipality dataset.

[4]All area computations were performed on an Eckert VI equal area projection of the municipality dataset in order to achieve correct area estimates.

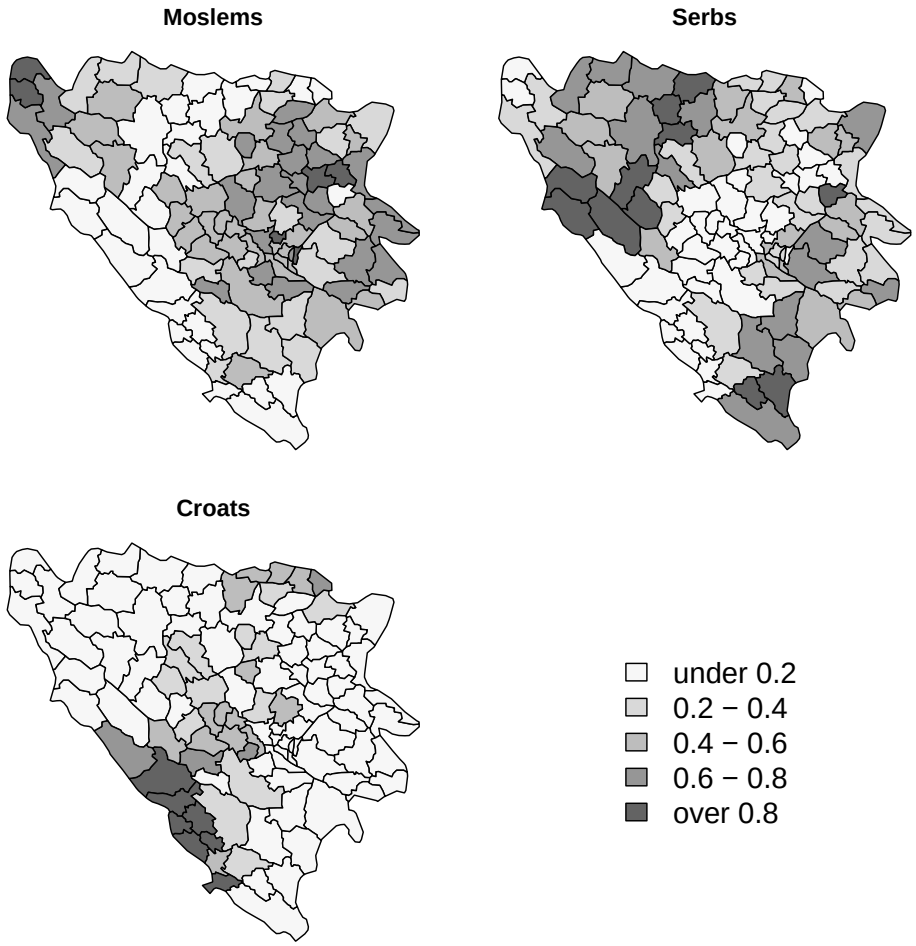


Figure 3.8: Map of the settlement areas of the three major ethnic groups in Bosnia, according to the 1991 census. For each of the ethnic groups, the plot shows the population shares at the municipality level.

3.5 Chapter Summary

In this chapter, I introduced the spatial data on settlement patterns used in this dissertation. Recent progress in the computational processing of spatial data makes it beneficial to use GIS software for the analysis in the subsequent chapters. GIS data comes in two fundamentally different formats – raster and vector data. Vector data represent discrete geographic features and can consist of either points, lines or polygons. The latter were used in the GREG dataset to represent the settlement regions of ethnic groups. GREG was created for this dissertation and contains a worldwide list of settlement areas of ethnic groups, based on the Soviet *Atlas Narodov Mira*. I discussed the creation of the GREG dataset and possible difficulties associated with it. An alternative way of capturing the geographic settlement patterns of groups is to utilize

census data. Some countries include the ethnicity of the respondent in their census, which can be used to infer the geographic settlement patterns of groups by mapping the census data on the administrative units the population figures are available for. I have presented a dataset on Bosnia that uses this approach to achieve fine-grained spatial data on ethnicity. Having outlined the data foundations for this dissertation, I now turn to the analysis of settlement patterns and conflict in the subsequent chapters.

Chapter 4

Spatial Concentration and Conflict Risk

In the theoretical introduction to this dissertation, I argued that settlement patterns can shape the opportunities for social interaction within a group and thus influence the risk of conflict. In following up recent work on group concentration and conflict, this chapter presents a first empirical test of this argument. There is consensus in the literature that geographic group concentration is related to a higher risk of conflict, but so far the literature has not been able to find out why this is the case. Two competing mechanisms have been proposed. The first is what I call "group concentration as motivation": Concentrated groups are more likely to see their territory as a homeland and thus more willing to fight for it. The second mechanism is an "opportunity" mechanism, and it corresponds to what I have outlined in Chapter 2: Concentrated groups can more easily overcome the collective action problem and can thus successfully mobilize for violence. Using the GREG dataset presented in the previous chapter, I show how we can distinguish empirically between the two mechanisms and test their independent effect on conflict risk. My results suggest that the "opportunity" mechanism is the driving force that links concentration to conflict. This chapter starts with a short review of the causal mechanisms linking concentration and conflict and proceeds to a discussion on how to measure them using the GREG dataset. I introduce the research design of the quantitative study and present the results. The chapter ends with an interpretation of the results and a summary.

4.1 Introduction

The literature on the causes on civil war shows a significant divide. Earlier work tends to focus on the actors' motivations for fighting: Conflict occurs because some groups in a society perceive themselves to be politically or economically disadvantaged compared to others (Gurr 1970; Horowitz 1985). According to this perspective, collective grievances are the main factor

that drives conflict. More recently, however, scholars have advanced a different perspective. Current work on civil war (Fearon and Laitin 2003; Collier and Hoeffler 2004) highlights opportunities as driving forces behind conflict: Groups fight because under certain circumstances, violence becomes a viable strategy to achieve material benefits. Whereas motivational explanations see conflict as the result of group discrimination, the thrust of the opportunity perspective lies in the context that makes violence profitable.

The 'motivation versus opportunity' (or 'grievance versus greed') debate continues to transcend different sub-fields of civil war studies. This chapter investigates the effect of group geography on conflict and distinguishes between geography as a motivation, and geography as an opportunity for violence. The question of how – and not just whether – group geography affects conflict is important: Not only do a significant proportion of intrastate conflicts involve territorial issues (Buhaug 2006), but territorial conflicts have also proven to be much more difficult to resolve (Fearon 2004). Obviously, group geography seems to be an important factor in bringing about violence. And still, our knowledge about the mechanisms linking group geography and conflict remains incomplete, which calls for a closer examination of the problem.

What can we learn from studying the impact of group geography on conflict? This chapter focuses on a particular aspect of a group's geography: geographic concentration. A group is geographically concentrated if it settles in one contiguous area of a country; by contrast, a fragmented group has various settlement areas scattered across the country. The study of group settlement patterns was pioneered by Toft (2002; 2003) who examined the relationship between geographic concentration and the level of violence a group faces. The conclusion of Toft's studies is that geographically concentrated groups face a higher risk of conflict, a proposition that is shared by other authors (Cornell 2002; Lichbach 1995; Goemans 2006).

How can we explain the impact of group concentration on violence? There are different explanations linking concentration to conflict. Broadly speaking, following Toft (2002, 2003) we can distinguish between a motivation- and an opportunity-driven mechanism. The former relates concentration of the settlement pattern to a group's motivation to take arms. The argument is that groups with a higher spatial concentration are more likely to show a strong attachment to their territory and should therefore be more likely to fight for it. The latter mechanism posits a relationship between the settlement pattern and a group's capabilities for fighting. More precisely, because of the facilitated interaction of group members, concentrated groups should be more likely to overcome collective action problems and therefore have a higher probability of engaging in conflict.

However, until now empirical analyses have been unable to distinguish between the two mechanisms, and we are left unsure of their relative importance. In this chapter I show that with

more advanced data and methods, we can overcome the limitations of previous work and test the two mechanisms independently. Recent progress in the application of Geographic Information Systems (GIS) in social science research allows for an increasingly detailed perspective on previously aggregated concepts, such as states or groups. When studying the geography of ethnic groups, GIS enables us to open the "black box" of a group's geography, and study more closely the intra-group processes leading to group conflict. Using GIS, I develop indicators measuring two different aspects of a settlement pattern: first, the degree to which the territory is concentrated or fragmented, and second, the degree to which the group population is concentrated or spatially separated. I argue that the first measure, territorial concentration, is what affects a group's motivation to fight, whereas the second indicator, population dispersion, affects the group capabilities and therefore captures an opportunity-driven conflict mechanism. In my empirical test, I find evidence in favor of an explanation that links conflict to facilitated interaction opportunities in concentrated groups, rather than a motivational mechanism.

4.2 Group Concentration as Motivation and Opportunity

The distinction between geography as a motivation or opportunity for conflict is not new. Geography and conflict have received extensive treatment in the literature on interstate conflict, and two broad perspectives can be identified: territory as a source of conflict, and as a facilitating condition of conflict (Diehl 1991). The first approach sees geography as a source of conflict because territory is an indivisible issue, which is why disputes over territory are likely to escalate into violent conflict (Fearon 1995). Scholars of international conflict commonly refer to three reasons why territory is valuable to states (Hensel 2000): First, territory has an objective value due to valuable resources located there. Second, territory has a strategic value because control over it gives a military advantage; third, territory has a subjective value because it is part of a state's culture and identity. Empirical analyses of interstate wars have shown that territorial issues are among the most frequent and robust determinants of war (Holsti 1991; Hensel 2000).

The second perspective on the role of geography in interstate conflict highlights the opportunities for war that are created by space. For example, it has been argued that geography determines the military reach of states, in that more proximate states can more easily inflict damage on each other (Boulding 1962; Lemke 1995). An alternative explanation suggests that immediate neighbors are more likely to be perceived as a threat to a state because there is less uncertainty regarding the future time horizon of the interaction (Starr and Most 1978). Yet another argument relates geographic proximity to a higher frequency of interaction, and therefore to a higher probability of conflict. Here, rather than being an issue of dispute, geography

determines the environment – the topology – in which territorial states interact.

How does geography operate in civil wars? With some limitations, the insights gained from interstate wars also apply to intrastate conflict. Geography can be the motivation for fighting. For example, the territory under dispute in civil war may contain valuable resources such as diamonds (Lujala, Gleditsch and Gilmore 2005) or oil (Lujala, Rød and Thieme 2007). Alternatively, territory may have strategic importance to the actors in civil wars: For example, controlling areas close to an international border secures cross-border supply of resources for fighting (Buhaug and Rød 2006). Lastly, there is also an intrinsic value of territory to a group (Toft 2003), an issue to which I will return later. Similar to interstate conflict, geography can also provide opportunities for fighting in civil war. Some studies have found support for an intrastate loss-of-strength gradient of state power: Groups and regions which are further away from the capital of a state tend to have a higher likelihood of being affected by conflict (Buhaug and Rød 2006; Buhaug, Cederman and Rød 2008). Similarly, rough terrain has been hypothesized to provide shelter for rebel groups from government forces, making mountainous regions more susceptible to conflict (Fearon and Laitin 2003; Buhaug and Rød 2006; Buhaug, Cederman and Rød 2008).

Group concentration is a geographic aspect that deserves particular attention in the study of civil war. A group is concentrated if its settlement area consists of a single, coherent cluster. Conversely, a group is fragmented if its territory consists of a set of smaller, detached areas. Group concentration has frequently been emphasized as a determinant of conflict (Horowitz 1985; Posen 1993*a*; Fearon 1998; Cornell 2002), but it was not until the work of Toft (2002; 2003) that it received systematic scrutiny. Toft's work shows that geographically concentrated groups exhibit higher levels of violence against the state. She outlines both a motivation- and an opportunity-driven mechanism to explain this. On the one hand, concentrated groups are more likely to see their territory as their homeland, and thus might be more willing to fight for it. On the other hand, concentrated groups face fewer difficulties in overcoming the collective action problem, and might therefore be more likely to successfully mobilize for conflict. Although Toft's work shows a positive relationship between concentration and conflict, it does not test the two hypothesized mechanisms independently. Essentially, we know that concentration is related to conflict, but not *why* this is the case. However, it is crucial to understand how and not merely whether concentration drives conflict: Settlement pattern adjustments, particularly the geographic separation of warring groups, have been proposed as solutions to ethnic civil wars (Kaufmann 1996). Without precise knowledge about the mechanisms linking group geography and conflict, we are unable to assess the expected efficacy of such measures: While territorial separation might terminate conflict in the short run, the artificial territorial clusters created by

those adjustments might also incite future war.

In the following paragraphs, I examine the mechanisms between concentration and conflict more closely, and link them to observable properties of a group's settlement pattern. In order to do so, I argue that it is necessary to refine the definition of group concentration. Previous research has treated concentration as a dichotomous attribute along a single dimension, assuming that a group's settlement pattern is located somewhere between perfect concentration and complete fragmentation. However, as it will become more apparent below, I argue that it is necessary to keep two aspects of concentration separate. First, concentration can be defined with respect to group territory: Do we find more than one area where a group settles, and if so, how large are the different settlement areas? This first dimension should link territory to group motivation for fighting: Efforts to establish and control a clearly demarcated group territory should obviously find much more resonance in the group population if the group's territory is clearly defined, which is not the case for groups with many territorial clusters across a state. Obviously, this first dimension focuses on territory only, and omits any aspect of the spatial distribution of the group population. An alternative definition of concentration is based on population: How clustered is the group population in space, or equivalently, how far are people of a group located from each other? This second dimension relates to the opportunity mechanism that links group concentration to conflict: As I described in Chapter 2, social mobilization should be less difficult if people are located close to each other. I proceed by examining the effects of territorial concentration and population dispersion on conflict.

4.2.1 Territorial Concentration and the Motivation for Conflict

As introduced above, in interstate conflict territory can incite conflict because it is likely to be seen as an indivisible issue (Fearon 1995). Toft's (2003) "Theory of Indivisible Territory and Ethnic War" applies this logic to intrastate conflicts. She argues that conflict arises if (i) the group demands control over its territory, which (ii) the state is not willing to grant. A territorially concentrated settlement pattern is likely to influence the first condition. Only if the territory of a group is reasonably well defined can we assume that the group considers it as "theirs". If, on the other hand, a group occupies more than one territory across a country, there is no clearly demarcated territory for which they could demand autonomy. Here, the value of a concentrated territory is twofold: First, a concentrated territory is more likely to be seen as a homeland for a group. For an ethnic group, its homeland is a crucial part of its identity. In the previous chapter, I showed that scholars commonly see territorial attachment as an integral dimension along which to distinguish ethnic groups. Besides language and culture, the ethnic homeland is one of the observable markers that helps the group develop an internal sense of belonging.

When making references to their homeland, groups typically give it a name. For example, Toft (2003) mentions the example of the Welsh, a group which is strongly attached to a particular homeland, Wales. However, for this to be possible, it is necessary that the group territory be concentrated. This is a consequence of the difficulty of labeling a non-contiguous piece of land with a common name: It is unlikely that a particular homeland would be considered as such if it was distributed over the area of a state in separate territorial chunks. The second value of a concentrated territory is strategic. If groups fight with the intention of gaining autonomy over their territories, the shape of their territory partly determines whether their efforts will be successful in the long run. A contiguous piece of land is much easier to defend and secure from possible future external threats (Posen 1993*a*). Also, as Goemans (2006) argues, the territorial concentration of a group helps it to overcome coordination problems within the group, as required for establishing a collective defense, for example.

The motivational mechanism linking territorial concentration and conflict then works as follows. The escalation of violence up to the level of ethnic war can only occur if there is mass support in the group population for conflict. However, for mass mobilization to be possible, the group population must believe that the cause for fighting is justified. This condition is more likely to hold for concentrated groups. As outlined above, a concentrated group attributes (i) a higher value to its territory and (ii) a higher likelihood to the feasibility of creating a new state from it. A concentrated group will then be more likely to mobilize for conflict, because the perceived cause for conflict – the control of its territory – will much easier resonate in the group population than in the case of spatially dispersed groups. In summary, higher degrees of territorial concentration are related to a higher likelihood of autonomy claims by a group. According to the above theoretical considerations, we should observe that

> H1: Higher territorial concentration of a group is associated with a higher risk of conflict involvement.

Having discussed how territorial concentration can provide a motivation for conflict, I now turn to the second aspect: the relation between population concentration and conflict.

4.2.2 Population Dispersion and Group Capability for Fighting

Group geography provides the context in which group members interact, and is therefore crucial in determining the group's ability to mobilize for conflict. In order to compete against government forces, groups need both the support and participation of their population, as well as material resources for fighting. Toft (2003) argues that concentrated groups have better political, economic and social networks that they can use to successfully start and sustain fighting.

Especially when it comes to the mobilization of the masses for conflict, geographic concentration and therefore spatial proximity of individuals proves advantageous. Other scholars have argued along similar lines. In his analysis of rebellion as a collective action problem, Lichbach (1995) emphasizes geographic proximity as an important factor which fosters coordination. He attributes the conflict-fostering effect of concentration to three factors: First, spatial proximity of group members leads to frequent interactions and to the formation of a "cognitive proximity" (p. 159) in the group. This cognitive proximity is especially important since it fosters the establishment of collective grievances among the group population. Second, concentration of the population makes it easier to coordinate collective action by reducing organizational costs. Third, concentrated groups can more easily sustain dissent by administering rewards and monitoring defectors (see e.g. Hardin 1995).

Essentially, the effect of geographic concentration on conflict comes about because people are located close to each other and can interact more easily than individuals who live further from each other. For that reason, what matters here is the extent to which the settlement pattern of a group allows for the interaction of people. Unlike in the previous section where I argued that territorial concentration provides a motivation for conflict, here we must study the location of people relative to each other to find out about a group's opportunities for mobilization. Furthermore, our measure of population dispersion must take into account the distance between people: The more distant the group members are located from each other, the less likely the group is to mobilize a critical number of people and resources for conflict. The geographic distance of group members is what I refer to as population dispersion, whereby highly dispersed groups should be less susceptible for conflict. Therefore, I hypothesize that

H2: Population dispersion of a group leads to a lower risk of conflict involvement.

The following sections describe the empirical analysis of the theoretical propositions introduced above.

4.3 Measuring Features of the Settlement Pattern

I now turn to the operationalization of my two main independent variables – territorial concentration and population dispersion. Both measures rely on the generalized GREG data as described in Chapter 3. Remember that in GREG, each group is digitally represented by a set of one or more polygons. For the purpose of illustrating the settlement pattern measures, Figure 4.1 shows an example of the data contained in the GIS dataset. The figure shows the distribution of the Ossetians in the Georgia. According to GREG, the Ossetians occupy one main territorial

cluster bordering Russia in the north. This cluster roughly corresponds to the region of South Ossetia. However, there are also much smaller Ossetian settlements distributed in the south and the east. Note that the large territorial cluster in the north is represented by more than one polygon (black lines). This is due to the fact that GREG distinguishes between mono-ethnic and mixed regions, each of which are represented as different polygons. For example, the largest polygon of the main Ossetian cluster in the north of Georgia denotes an exclusively Ossetian area, whereas the smaller polygons at the lower part of this cluster correspond to mixed regions that the Ossetians share with another group, the Georgians.

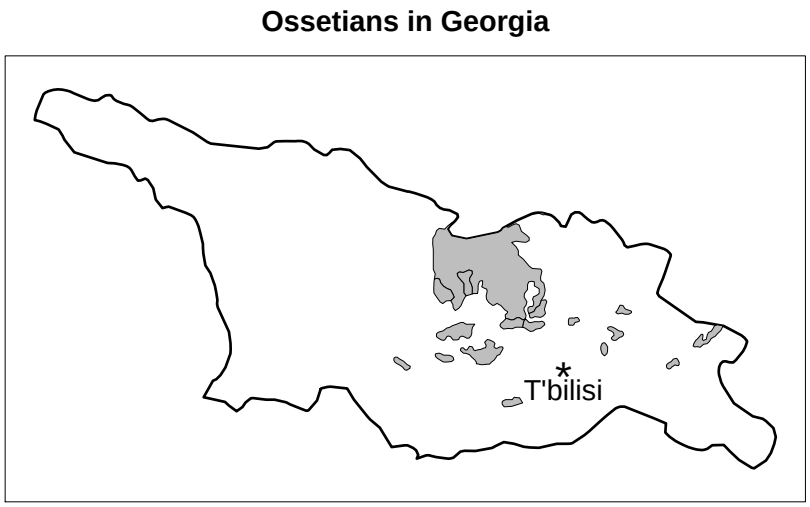


Figure 4.1: Example: GIS representation of the Ossetians in Georgia.

Based on the spatial data from GREG as shown in Figure 4.1, for each group I compute measures for territorial concentration and population dispersion needed for this analysis. Both measures are based on the set of territorial clusters among the group polygons. A cluster c_i is a set of contiguous polygons: For example, the main territorial cluster of the Ossetians in Figure 4.1 consists of six smaller polygons. Given the N clusters, I also require area and population estimates for each of them. The area of a cluster can be computed using a GIS.[1] Population estimates are obtained by overlaying the GREG polygons with a raster dataset on population [2]. Both area $a(c_i)$ and population $p(c_i)$ of a cluster c_i are proportions of the total group area or population, respectively.

[1] Area computations require a transformation of the polygons into an equal-area projection (here, Eckert VI).

[2] In an overlay procedure, the population cells covered by a polygon are selected and the sum of their population values is computed. The population raster used for this analysis is the *Gridded Population of the World* dataset released by CIESIN (2005).

Obviously, the settlement pattern of the Ossetians shown in Figure 4.1 is much more concentrated than that of a group with the same number of cells, but spread out in three equal-sized clusters. Ideally, the concentration measure should take into account not only the number but also the sizes of the different clusters. The Herfindahl concentration formula does exactly this. It is frequently used in economics to describe the degree of market concentration, but has also been applied to e.g. the ethnic composition of a country known as the ethno-linguistic fractionalization (ELF) index (Taylor and Hudson 1972; Alesina et al. 2003). The computation of a territorial concentration measure for the spatial distribution of an ethnic group is straightforward and works according to the following formula:

$$\text{Territorial concentration} = \sum_{i=1}^{n} a(c_i)^2$$

where the $a(c_i)$ is the area of cluster c_i. Territorial concentration takes a value in the interval $(0, 1]$ where low values correspond to scattered groups, and high values to territorially concentrated ones.

A measure of population dispersion is somewhat more difficult to obtain because one needs to take into account the distances between clusters. The index must be constructed in such a way that the distance between heavily populated clusters affects the final value of the index more than the distance between two sparsely populated clusters. This is achieved by computing a population-weighted average of the logged minimal geodesic distances between clusters according to the following formula:[3]

$$\text{Population dispersion} = \frac{\sum_{i<j\in N} p(c_i)p(c_j)log(d(c_i, c_j))}{\sum_{i<j\in N} p(c_i)p(c_j)}$$

where the c_i are the clusters of the group, $p(c_i)$ is the population of a cluster, and $d(c_i, c_j)$ is the geodesic minimal distance between two clusters.

4.4 Research Design and Data

In order to test the impact of territorial concentration and population dispersion on ethnic conflict empirically, I employ a time-series cross-sectional design, relying on a dataset created by Buhaug, Cederman and Rød (2008). The units of analysis in this dataset are politically

[3]The minimal geodesic distance between two clusters is the distance between a pair of cells, one from each cluster, such that the distance is minimal across all pairs.

marginalized ethnic groups – in other words, groups that do not participate in the government of a country and would therefore have a reason to resort to violent means. These groups are observed annually, covering the post-World War II period until 2003. The dataset presently covers 72 countries in Europe, Asia and North Africa.

The dependent variable is the onset of ethnic conflict in the post-World War II period between the respective marginalized ethnic group and the government, and is taken from Buhaug, Cederman and Rød (2008). More precisely, the authors code a conflict onset for a given group-year if (i) there was a civil war onset according to the *UCDP/PRIO Armed Conflict Dataset* (Gleditsch et al. 2002), and (ii) if the war was fought along ethnic lines, with the respective group involved. The latter information was supplemented from various external sources (refer to Buhaug et al. for details).

The main independent variables are territorial concentration and population dispersion, as described in the previous section. By definition, territorial concentration takes values in the $(0, 1]$ interval. In the sample, the variable ranges from 0.01 to 1, the latter corresponding to perfectly concentrated groups with only one territorial cluster. In line with H1, more concentrated groups should have a higher risk of conflict, so this variable is expected to have a positive effect. Population dispersion takes positive values and has a sample maximum of 8.4 (note that as described above, distances are log-transformed for the computation of dispersion). As I have argued above, more dispersed groups should be less prone to conflict, so I expect this variable to have a negative effect.

I add both group-level and country-level controls. In order to distinguish the group concentration effect found in earlier studies from my more detailed indicators, I include a dummy variable for concentrated groups, indicating whether the group settlement pattern consists of exactly one territorial cluster, as defined above. This variable provides a rough distinction between concentrated and non-concentrated groups, but cannot give a more nuanced description of the group area or population, as I require for this analysis. Other group-level control variables include the power balance between the government and the ethnic group as introduced by Buhaug, Cederman and Rød (2008). Power balance is expected to have a positive effect, because increasing strength compared to the government should make a group more likely to engage in conflict. I add two further control variables for the group's geography, both also taken from Buhaug, Cederman and Rød (2008): a measure of the group's distance from the capital, and an indicator for mountainous group terrain. For the former, a positive effect is expected since we can assume that more distant groups are more difficult to repress. By the same token, I also expect a positive effect for the latter variable, since we can assume that groups in the mountains escape state control more easily. At the country level, I control for the factors

that are commonly associated with civil war. Economic performance has been found to be a robust and highly significant predictor of civil war (Fearon and Laitin 2003; Collier and Hoeffler 2004), with poor states experiencing significantly more civil wars than wealthy ones. I use GDP per capita as an indicator for economic conditions and expect a negative effect on conflict. Also, larger countries tend to have more civil wars, so I include the logged country population in the regression. Regime type should also affect conflict propensity. We can assume that democratic countries have institutions to solve internal disputes peacefully and should therefore see less conflict. Consequently, I include a binary regime type indicator from Polity IV (Marshall and Jaggers 2000), coding a country as a democracy if it has a polity score of 6 or above. This variable should have a negative effect. Lastly, I control for the trend of increasing civil war propensity over time by including the year of observation in the regression. Complete descriptive statistics of all variables are given in Table 4.1.

Variable	Mean	Std.dev.	Min	Max
Conflict	0.003	0.056	0	1
Territorial concentration	0.739	0.3	0.017	1
Population dispersion	2.343	2.226	0	8.4
Power balance (log)	-5.723	2.227	-14.808	-0.323
Capital distance (log)	6.39	1.021	2.193	8.834
Mountains	0.536	0.362	0	1
GDP (logged)	7.676	0.865	5.713	10.206
Country population (log)	10.722	1.729	6.351	14.029
Democracy	0.258	0.437	0	1

Table 4.1: Summary statistics.

4.5 Results

4.5.1 Logit Regression Models

I estimate logit models with conflict onset as the dependent variable, and standard errors clustered by country.[4] Since observations are likely to show strong time dependence, I employ the method proposed by Beck, Katz and Tucker (1998) and include the number of peace years as well as three natural cubic splines. Table 4.2 shows the results of the regression analysis.

In Model 1, I start by replicating the finding from earlier research on settlement patterns, namely that concentrated groups have a higher likelihood of conflict. The results provide strong

[4]Software: StataSE 9.2.

	(1)	(2)	(3)	(4)	(5)
Territorial concentration		0.760		-0.0144	0.0631
		(0.568)		(0.641)	(0.628)
Population dispersion			-0.362*	-0.364*	-0.364*
			(0.158)	(0.168)	(0.163)
Concentrated group	0.487*	0.202	-0.946	-0.948	
	(0.225)	(0.328)	(0.574)	(0.557)	
Power balance	0.668**	0.661**	0.700**	0.700**	0.808**
	(0.109)	(0.110)	(0.128)	(0.128)	(0.132)
Capital distance	0.602**	0.575**	0.647**	0.648**	0.568*
	(0.214)	(0.216)	(0.200)	(0.205)	(0.268)
Mountains	1.291**	1.248**	1.182**	1.183**	1.566**
	(0.304)	(0.291)	(0.310)	(0.305)	(0.531)
GDP p.c.	-0.212	-0.198	-0.145	-0.144	-0.106
	(0.238)	(0.238)	(0.241)	(0.240)	(0.270)
Country population	0.315*	0.313*	0.316*	0.316*	0.372*
	(0.125)	(0.123)	(0.127)	(0.127)	(0.154)
Democracy	0.753**	0.747**	0.719**	0.719**	0.899*
	(0.218)	(0.211)	(0.222)	(0.225)	(0.378)
Year	0.0648**	0.0643**	0.0638**	0.0638**	0.0511**
	(0.0124)	(0.0125)	(0.0122)	(0.0123)	(0.0139)
Peace years	-0.248**	-0.249**	-0.245**	-0.245**	-0.229*
	(0.0783)	(0.0784)	(0.0771)	(0.0767)	(0.107)
Constant	-135.4**	-134.8**	-132.5**	-132.5**	-107.6**
	(23.36)	(23.39)	(23.14)	(23.13)	(26.38)
Observations	26853	26853	26853	26853	15072
Log-pseudolikelihood	-460.741	-459.678	-456.661	-456.661	-254.12

Table 4.2: Logit regression results, with robust standard errors in parentheses. $^{**}p < 0.01$, $^{*}p < 0.05$. Results for the splines are not shown.

evidence for this: The coefficient for the group concentration dummy is positive and significant. The control variables turn out largely as expected. In line with Buhaug et al.'s (2008) results, we see that power balance in favor of the group, distance to the capital, and mountainous group territory all make conflict more likely. Economic performance has no discernible relationship to conflict. However, groups in larger countries face a higher conflict risk, and so do groups in more democratic countries. The latter effect runs counter to my expectations, but might be due to the fact that democratic governments face more constraints in using repression to avoid civil unrest. We also see a clear trend of increasing conflict propensity over time, as shown by

the positive and significant coefficient of the year of observation variable. Also, there is a clear time dependence of conflict observations in that conflict tends to become less likely the longer a group remains at peace. This is indicated by the negative and significant sign of the peace years variable.

Models 2 and 3 add my variables for territorial concentration and population dispersion, one at a time. The results from Model 2 provide no evidence that my more nuanced measure for concentration has any explanatory power: The territorial concentration variable receives a positive coefficient, but is far from significant. However, the picture is different in Model 3. Population dispersion seems to be a strong predictor of conflict, and clearly outperforms the simple concentration dummy. The effect of population dispersion is strong: An increase in dispersion from the 10th to the 90th percentile reduces the predicted probability of conflict from 0.2 percent to 0.03 percent, while keeping the other variables at their means.

Model 4 performs a joint test of my indicators for territorial concentration and population dispersion. This model confirms the result obtained in the previous one: Population dispersion again comes out as negative and significant, whereas territorial concentration has no effect. However, one more test is required before we can reach a final conclusion. For perfectly concentrated groups (i.e. groups settling in a single cluster), territorial concentration and population dispersion are perfectly correlated. This is due to the fact that for these groups, territorial concentration always takes the highest value of 1, whereas population dispersion is 0. As a result, in Model 4 these variables have a high degree of correlation (-0.79), thus potentially giving rise to a multicollinearity problem. For this reason, I re-estimate Model 4 on a restricted sample that only includes only groups with more than one territorial cluster. This reduces the sample size to slightly more than half of the original one, but with territorial concentration and population correlated only at -0.35. The results are reported as Model 5 in Table 4.2. Again, we see the same result as in the previous models. Population dispersion has a negative effect on conflict, but territorial concentration remains insignificant. The marginal effect of population dispersion in Model 5 amouts to a decrease in the predicted conflict probability from 0.12 percent to 0.04 percent, while keeping the other covariates at their means. In sum, these initial results from the regression analysis show a clear picture: Obviously, what drives the relationship between concentration and conflict is the dispersion of the group's population, rather than the division of the territory into multiple clusters. In other words, the empirical evidence supports an opportunity-driven mechanism from group concentration to conflict.

4.5.2 Robustness Checks

In order to determine whether the results obtained above are generally valid, I conduct a series of robustness checks. First, I check whether the group definition criteria employed by the Atlas Narodov Mira affect the result. The Atlas employs a fine-grained group distinction, so there is a risk of including a number of small groups which have little political relevance. Therefore, I repeat Models 4 and 5 on a restricted sample, including groups with a dyadic power share of 0.01 or more, as coded by Buhaug, Cederman and Rød (2008). This procedure follows the robustness checks provided in the original study, and leads to a restricted sample containing 8,214 group-years from 57 countries. The results obtained from this sample again provide strong evidence for the effect of population dispersion. Both Models 4 and 5 estimated on this sample yield a negative coefficient for population dispersion which is significant at the 0.01 level.

My second robustness check repeats the analysis in a cross-sectional design, where groups in the post-World War II period constitute the units of analysis. Besides my group concentration variables, this analysis includes the same group level variables as described above (power balance, capital distance, mountains). At the country level, I control for economic performance and democracy, both of which are based on the first available year for the respective country – that is, 1946 or the year the state entered the international system. The sample consists of 750 groups from 69 countries (and 436 groups from 61 countries when restricting it to non-concentrated groups, as required for Model 5). Again, an estimation of Models 4 and 5 on the cross-sectional data support the conclusions reached above. In both models, population dispersion receives a negative and significant coefficient, whereas territorial concentration has no effect.

4.6 Chapter Summary

What can we learn from the empirical results presented in this chapter? Most importantly, I have shown *how* group concentration affects conflict. The scholarly literature has so far been unable to explain which mechanisms drive the causal link from group geography to conflict. When comparing the motivation and opportunity mechanisms linking group concentration and conflict, my results point to the relative importance of the latter. In other words, what seems to matter for explaining a group's conflict propensity is an opportunity-driven mechanism: the more proximate a group's members, the better the group is able to coordinate for collective action. The distinction between the two mechanisms is made possible by the development of new measures of ethnic settlement patterns. Using the GREG GIS dataset on ethnic groups introduced in Chapter 3, I developed indicators that are able to distinguish between group geography

as motivation or opportunity, something existing measures cannot capture.

Even though this analysis went one step further than existing work, there is still need to study the mechanisms linking geography and conflict more precisely. For example, in order to specify a causal mechanism linking group geography to better group interaction opportunities and ultimately to conflict, we would have to specify the actors involved in mobilization and describe precisely how geography affects the mobilization process (Hedström and Swedberg 1998). However, although this chapter fails to make the precise mechanisms explicit, it points to the importance of social interaction and collective mobilization within groups. Consequently, I pick up these issues in the next chapter, where I provide a more detailed discussion of the group mobilization process and the impact of group geography by means of computational modeling.

Chapter 5

Simulating a Spatial Mechanism of Group Mobilization

SPATIAL concentration matters, primarily as a facilitating condition for group mobilization – this is what I have shown in the previous chapter. However, even though the more advanced spatial indicators help us get closer to the hypothesized mechanisms within a standard regression framework, they still constitute an indirect test: Ultimately, we do not know whether the conflict proneness of concentrated groups is solely due to the facilitated interaction of group individuals, since the empirical test omits actors and interaction processes that are integral to the proposed relationship. Thus, what is the right way to test this argument? The answer to this question lies at the core of a great methodological divide in political science. On the one hand, using case studies, we would be able to produce rich and detailed insights into group processes during conflict (see e.g. Wood 2003; Kalyvas 2006). However, these approaches typically focus on a small number of cases, making it difficult to generalize findings to a larger sample. On the other hand, quantitative approaches as I adopted in the previous chapter, identify patterns and regularities in the occurrence of ethnic conflict, but make it difficult to study precisely the mechanisms driving these relationships.

In order to bridge this divide, in this chapter I take an intermediate approach and develop a computational model of group mobilization. This model makes the proposed mechanisms of group mobilization outlined in the previous chapter explicit by translating them into an executable computer program. Essentially, the model contains a simple representation of the social actors and simulates their interaction in a geographic space, built on real-world GIS data. The approach of linking models to real geographies is new and has rarely been applied to social science questions (Gimblett 2002). The model serves as an "operational" theory – a set of complex interdependencies that can be used to study how my independent variable, group settlement

patterns, affects group mobilization in a real set of cases. Using computational modeling, we therefore move from pure correlational studies to a richer – but also more complex – test of the proposed argument.

The use of computer models in the social sciences is not a new trend, but certainly goes beyond mainstream methodological approaches in the field (Cederman 2001). For that reason, I begin this chapter with a review of computer modeling and its theoretical backgrounds. Following this introduction, the bulk of this chapter is devoted to my computational model of group geography and mobilization. Using GIS data on the spatial distribution of ethnic groups, I simulate the spread of mobilization between group members. For each group, the model computes the degree to which the hypothesized mobilization process affects the population. Using regression analysis, I then test whether thorough mobilization according to the model is related to a higher conflict risk, as predicted by my theory.

5.1 Models as Mediators between Theory and Data

I start with an overview of different approaches to scientific explanation and show how models can serve as supplementary mediators between a theory and its empirical realization.

5.1.1 Causal Explanations in the Social Sciences

Unlike in the natural sciences, the concept of causality is oftentimes difficult to pinpoint in the social sciences. What constitutes a causal explanation for a particular phenomenon? The predominant approach in the natural sciences is the *deductive-nomological* model of explanation (Little 1991). Here, a particular effect or event is considered explainable if it can be subsumed under a general covering law (Hempel 1965). For example, in order to explain why a ball dropped from a window falls to the ground, the general law of gravity is applied which states that objects with mass attract each other. Since the law is general, the effect will occur under any circumstances.

This claim of generality poses problems for the definition of a causal explanation in the social sciences. Rarely do social scientists deal with law-like relationships. In social systems, cause and effect are related to each other only across a series of intermediate steps, which in most cases exert independent influence on the relationship and therefore blur it. This is why instead of a perfect cause-effect relationship, we may fail to observe the effect in individual cases, and would therefore have to dismiss the supposed cause as an explanation under the deductive-nomological model. In the social sciences, the predominant model of explanation is *inductive-statistical* (Little 1991). Here, rather then requiring the effect to occur as a strict

consequence of the cause, the relationship is probabilistic. In other words, the presence of a given cause makes the effect *more likely*, but does not require it to occur under all circumstances. The analysis presented in the previous chapter serves as an example. I have shown that group concentration is related to a higher risk of conflict. Clearly, under the deductive-nomological model of explanation, concentration is not a sufficient cause of conflict, because there are a large number of concentrated groups that have remained peaceful. However, under the inductive-statistical model, group concentration constitutes an explanation for conflict because there is a significant probabilistic relationship.

However, the inductive-statistical model has come under some criticism within the social science community (see e.g. George and Bennett 2004, ch. 1). This is because statistical studies yield evidence of covariations of variables, but they cannot reveal the causal steps that brought about a particular effect. In the group concentration analysis presented in the previous chapter, the regression analysis shows that the distance of a group's population clusters is related to the occurrence of violence, but we cannot tell whether this effect is really due to group interaction opportunities, as the theory suggests. In short, with statistical approaches we can show that concentrated groups have a higher risk of conflict, but we are unable to explain *why*. Dissatisfied with inductive-statistical explanations, many scholars advocate what has been termed a *mechanistic* approach (Sawyer 2004) to explanation, aiming to fill the "black box" between an independent and dependent variable. A causal mechanism that links the two must "state *how*, by what intermediate steps, a certain outcome follows from a set of initial conditions" (Mayntz 2003, p. 241, emphasis original). It is important to mention that scholars advocating a mechanistic approach tend to require some generality for what they define as a mechanism (Hedström and Swedberg 1998). Unlike for historians, an individual realization of a chain of events alone does not constitute a mechanism. Rather, as Mayntz defines it, mechanisms are "causal generalizations about recurrent events" (Mayntz 2003, p. 241).

When explaining macro-level correlations – for example between group concentration and conflict risk – there is an additional complication we need to take into account. In most cases, the phenomena political scientists deal with are aggregate properties of societies, such as the spatial distribution of individuals and ethnic conflict discussed here. Simply put, researchers observe only the result of the interaction of many individuals. However, for a valid causal account of this outcome, we need to explain how these individuals interact to produce the observed aggregate outcome. The perspective of the individual as a causal agent is the "methodological individualism" inherent in the social sciences. Coleman (1990) summarizes the interaction between the macro-level of observed societal conditions, and the micro-level of interacting individuals. Figure 5.1 shows this as applied to the topic of this dissertation.

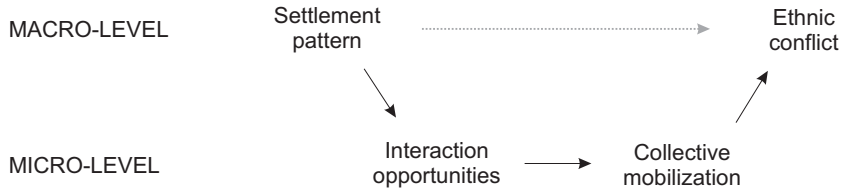


Figure 5.1: Methodological individualism: Explaining macro-level relationships with micro-level interactions.

At the macro-level, the relationship we are interested in is the effect of geographic group settlement patterns on ethnic conflict (shown as a grey arrow). In Chapter 2, I argued that this effect is brought about by the interaction opportunities of group members as enabled by geography. This mechanism can be seen as a sequence of three causal steps, shown as black arrows. In the language of Hedström and Swedberg (1998), the first arrow (left) is a *situational* mechanism – the spatial arrangement of individuals determines possible interaction opportunities. This allows for possible mobilization attempts to spread through the group population (second arrow, bottom) as group individuals choose to participate or abstain. This is what Hedström and Swedberg (1998) call an *action-formation* mechanism. Finally, in a *transformational mechanism*, the aggregate participation of many individuals eventually results in ethnic conflict (right arrow).

Scholars studying mechanisms usually advocate process-tracing as their empirical method of choice (Checkel 2008; George and Bennett 2004). In a process-tracing analysis, the hypothesized process is carefully scrutinized by examining the interactions between the intermediate causal agents. For example, group mobilization for conflict could be studied on a case of ethnic conflict, carefully tracing how unrest spread through the population, eventually reaching the level of full-fledged war. Process-tracing requires a detailed case analysis of the actors and events that are causally prior to the phenomenon in question. Few cases can be examined at such a high level of detail. Moreover, in process-tracing case studies it is difficult to establish causality: Would the outcome be the same if some event *A* had not occurred, or if some actor *B* had reacted differently? The advantage of process tracing is that the analysis is tightly linked to the causal agents and their actions. However, this richness necessarily comes at the expense of generality. This is why scholars have made the case for using computational modeling to study mechanisms (Cederman 1997). As compared to process tracing, computer models employ a higher level of abstraction, but allow the researcher to study systematic variations in the hypothesized processes. The next section gives a short overview of existing computational modeling approaches in the social sciences. I then move on to a more recent development, so-called agent-based models. The mobilization model I introduce below shares many features

with the agent-based approach, so it deserves a closer look.

5.1.2 Models in the Social Sciences

Generally speaking, a model is an abstract representation of particular aspects of reality that are relevant for the phenomenon in question (Gilbert and Troitzsch 2005; Gilbert 2007). For example, a toy model of a new plane helps us to study the aerodynamics of the prototype, without having to build the plane to scale. By definition, a model simplifies reality by leaving certain aspects out of the model, and making the included aspects of reality much simpler. In order to analyze the aerodynamics of the plane prototype, we do not require a precise implementation of the plane's internal infrastructure – all that matters is the outer shape of the plane, and this is where the model needs to resemble the original plane precisely. In general, models can be characterized by two features (Frigg and Hartmann 2006): First, their ontology – that is, how the relevant aspects of reality are represented. For example, the above-mentioned plane model can simply consist of a miniature representation of the aircraft. Second, models feature a way new knowledge is derived from them – their epistemology. The plane model, for example, can be analyzed in a wind tunnel in order to draw inferences on the aerodynamical properties of the real plane.

Modeling human societies is in some ways more difficult than modeling physical systems. Above all, a human society consists of large number of individuals, none of whom acts in simple and predictable ways. Social scientists can observe this complexity only partially. For example, in this dissertation, we observe group concentration and an increased risk of conflict, but not the steps in between. A model helps to relate these partial observations to each other by filling in the missing "gaps". Many social science models rely on a formal ontology where the relevant objects of study – individuals or organizations – are represented as mathematical entities. However, these formal models differ (i) with respect to the level at which this ontology operates (i.e. the societal level, or the individual), (ii) with respect to their epistemology (i.e. by solving the mathematical model analytically, or by means of simulation).

Early applications of models in the social sciences focused on aggregate quantities of social systems. Mostly developed by economists, these *analytical macro-models* used differential equations (for an overview, see Gilbert and Troitzsch 2005). However, more complex systems of interrelated differential equations are difficult and often impossible to solve analytically, so computer simulation was increasingly applied to derive model predictions. The *system dynamics* approach became popular through the work of Jay Forrester at MIT, who performed large-scale simulations at the global level, modeling the interactions of economic, political and environmental systems. His work lay the foundations for *Limits to Growth* (Meadows et al.

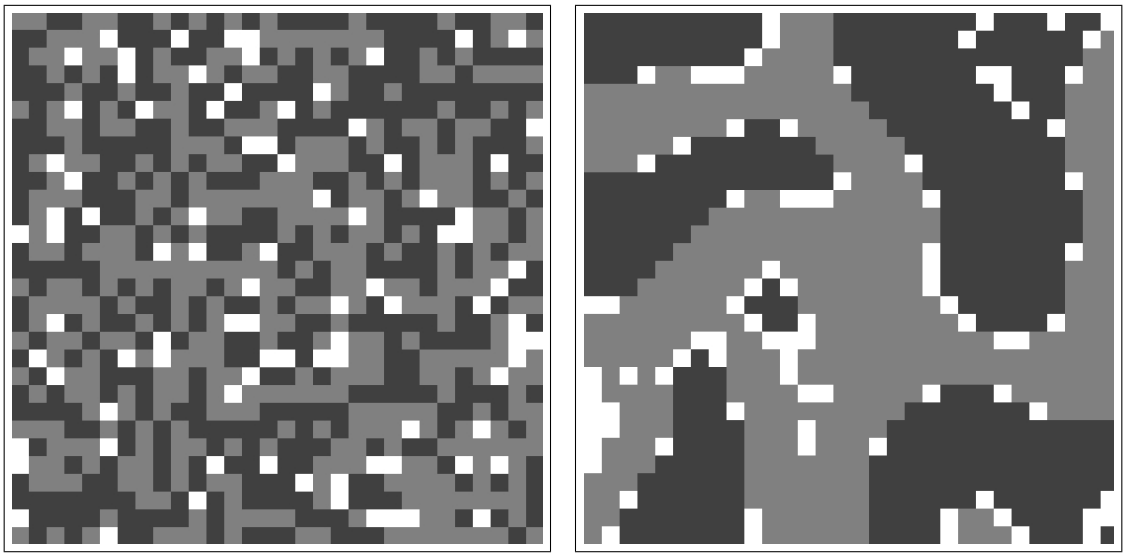
1972) – an influential book commissioned by the Club of Rome, which assessed economic and societal growth's consequences on the depletion of natural resources. The conclusions of the book were widely questioned, but system dynamics as a method continues to be applied, often using the STELLA software package (Gilbert and Troitzsch 2005).

System dynamics models operate exclusively at the macro-level and hide the inner workings of a society behind abstract equations. Game-theoretic models, in contrast, focus exclusively on social actors (Osborne and Rubinstein 1994). A game is a strategic setting in which a small number of decision makers interact. These actors are assumed to be rational individuals, i.e. they choose the action that maximizes an expected outcome. However, the decisions of these individuals are contingent on each other, and game theory provides the tools to derive the expected choices and the conditions under which they occur. This solution is found analytically. However, for this to be possible, game-theoretic models must make strong simplifying assumptions. For example, games usually involve only a small number of actors, all of which are assumed to be unitary (on this point, see also Chapter 2 for a discussion of game-theoretic models of ethnic conflict). Essentially, this means that when representing for example an ethnic group as a game-theoretic actor, the group is assumed to make decisions in unison and is thus treated as if it were a single agent. As Cederman (1997) argues, this assumption is often overly simplistic and makes game-theoretic models unsuitable for the study of many social science phenomena. This has led many scholars to employ less restrictive modeling approaches. These approaches are mainly characterized by giving up the analytical solution concept used in game theory in favor of a simulation approach. I turn to these agent-based models in the next section.

5.1.3 Agent-based Computational Models

Agent-based models (ABMs) are "artificial worlds populated by agents that interact in non-trivial ways" (Cederman 2001, p. 16). These models differ from the above-mentioned system dynamics approaches in that they focus on the individuals as causal agents. This is consistent with methodological individualism as a way of explaining aggregate social outcomes. At the same time, ABMs are less restrictive than game-theoretic models in their assumptions about the agents and their interactions, because they rely on computer simulation to derive the results of a model. Computer simulation is a particular modeling technique in which a system (for example, a human society) is implemented as a computer program, which then executes the behavior of the system over a number of time steps in order to generate certain outcomes of interest (Gilbert and Troitzsch 2005). Before turning to a discussion of agent models in general, I describe briefly an example of an ABM, since it nicely illustrates the core features of this method that set it apart from other modeling approaches.

One of the first "agent-based" models in the social sciences is the segregation model by Schelling (1971, 1978), developed to explain racial segregation in urban neighborhoods. In line with the agent-based paradigm, Schelling provides an explanation for a macro-phenomenon (the segregation of neighborhoods) based on the choices of individuals. Unlike many of today's ABMs, the original model was not run on a computer. Instead, Schelling used a sheet of grid paper, and randomly placed pennies and dimes in the cells. The coins represented agents of two different groups. These agents made movement choices dependent on the their immediate neighborhood: If an agent was surrounded by a sufficient number of agents of the same kind, it would feel comfortable and would not move. However, if a certain threshold of members of the other group lived near a particular agent, this would be a sufficient reason for this agent to move to another free location in the grid. During a simulation run, agents were picked at random and these then decided whether to move, as described above. From an initially mixed settlement pattern (Figure 5.2, left), a set of large clusters of equal agents emerged (Figure 5.2, right) – a segregated landscape. Given the setup of the model, it is not surprising that these clusters emerged eventually. However, Schelling's result showed that this happened even for surprisingly high levels of tolerance. In other words, people do not have to be extreme racists for urban segregation to occur.



(a) (b)

Figure 5.2: Screenshots of the Schelling segregation model at different times during a simulation run. Each agent (shown as a square) belongs to one of two groups (dark or light grey). Left: The pattern at the beginning of the simulation displays a high degree of mixing. Right: Over time, homogeneous clusters emerge.

Schelling's model has a number of features that are characteristic for agent-based models

(Epstein 2006, p. 6-7): Agents are *heterogenous*, i.e. they prefer different neighborhoods; they are *autonomous* in deciding where to move; they are located in an *explicit space*; they *interact locally*, i.e. only with their neighbors; and they have *limited rationality*.

Why should an ABM like the Schelling model be useful to study group mobilization? As the previous chapter has shown, there is some indication that group geography affects mobilization for collective action, but we are unsure as to how this works. I have presented different arguments that link group geography to interaction opportunities and group mobilization, but so far this mechanism remains only loosely specified. What we need is a more stringent representation of the hypothetical process which allows us to assess its logical validity, but also to explore its boundaries, i.e. the conditions under which the relationship breaks down (Hoffmann 2008). This can be achieved using agent-based modeling.

As the name suggests, ABMs focus on the actors in a society. When modeling group mobilization, we ideally want to have a computational representation of the group members involved. Obviously, for this to be possible, many agent-based modelers choose to include only a small group "population" in their model, because the representation of the entire group would not be computationally feasible (see e.g. Bhavnani 2006; Srbljinovic et al. 2003). In these models, however, empirical validation is more difficult since the artificial world created there needs to be matched to the observed reality (see below for a more detailed discussion of this validation process). This is particularly so in the context of geographic settlement patterns. If we claim the relationship between the spatial distribution of a group and conflict risk to be driven by interaction opportunities in the population, the best possible test in a computational model should rely on the real geography of a group, and not on an artificially created one. For that reason, the model introduced below takes as input the group geography according to the GREG dataset. In order to run a mobilization process on this real group geography, we must deviate slightly from the ABM approach such that no individual agents, but rather small groups of them are represented in the model. Whereas this gives up distinctive ABM features such as agent heterogeneity and limited rationality, my approach retains the notion of an explicit space, local interactions, and autonomy, according to Epstein (2006).

The approach of building a model based on theoretical premises and using it to explore the boundaries of the proposed theory is not unique to ABM. What is distinctive, however, is that the model derives its outcomes from a geo-referenced population of heterogenous groups of agents. This process of "growing" a particular macro-level phenomenon – such as ethnic conflict – from the interactions within an agent population has been termed a "generative" approach to social science (Epstein 2006). If the proposed model is able to generate the observed outcome, the implemented mechanism is a candidate explanation for the macro-level relationship we are

interested in. The remainder of this chapter applies this generative approach to the study of group mobilization and conflict.

5.2 Social Interaction and Group Mobilization

This and the following sections introduce a computational model that simulates the spread of mobilization through a group population. The purpose of the model is to examine how the spatial arrangement of a group population given by the settlement pattern affects the group's proneness to mobilization. The model operates on basic spatial units: the group cells. Starting with a few extremists located in some of these group cells, over time mobilization spreads to other cells. I assume that the proximity of two cells is positively related to the movement of people between them and should therefore facilitate the spread of revolutionary ideas and activity between them. The settlement pattern of a group – where people live, and how close they are to each other – can then facilitate interaction and therefore mobilization, but can also have a detrimental effect if the settlement pattern inhibits interaction. As a general tendency, we should then see a higher probability of conflict involvement for groups whose settlement pattern facilitates mobilization.

A best possible test of this proposed mechanism of mobilization "contagion" would require us to collect large amounts of data on observable mobilization events, both over space and time and also for different groups. I resort to a different way of assessing the empirical accuracy of the model: I model the hypothesized mobilization process on real-world data using computer simulation. The output of this model is a measure of the difficulty of group mobilization. In other words, it tells us how thoroughly mobilization would spread through the group population if the proposed mechanism was at work. What remains to be done is to compare this mobilization measure to the empirical outcome: Are groups with a high mobilization speed indeed more prone to conflict? This comparison is made statistically on a larger sample of groups, and is described in the empirical section below.

Computational and formal models of group mobilization already exist in the literature. Threshold models of collective behavior (Granovetter 1978; Kuran 1989, 1991; Chwe 1999) provide an account of how the mobilization of an entire group can develop from a few initial extremists. It is assumed that both benefits and costs of rebellion are assessed in terms of the total number of people that have already joined the movement. However, the assumption that people have complete information about the actual number of participants requires full connectedness across the population, which, as Granovetter (1978, p. 1431) himself mentions, is often not realistic. This criticism applies especially to ethnic groups that usually cover large areas.

Instead, an individual's face-to-face contacts and social networks will be crucial in affecting an individual's participation decision. In Chapter 2, I have outlined three mechanisms by which direct social interaction can foster the spread of mobilization from some individuals to others: First, close interaction facilitates the socialization of individuals to common beliefs and values; second, social interaction enables the flow of information about group action that is currently being taken; and third, direct contact between individuals imposes social pressure on people to participate.

The first effect – the social impact on individual motivation – has frequently been emphasized by scholars of collective action (see for example Tarrow 1994). People's immediate networks have a high influence on their opinion, especially regarding participation in a social movement. Social impact theory, developed in the 1980s primarily by the social psychologist Bibb LatanÈ (Latané 1981; Nowak, Szamrej and Latané 1990), claims that the more proximate (spatially) the peers who exert influence, the stronger the effect on an individual's opinion. This socializing effect is especially relevant in the context of ethnic mobilization: If a single group member feels discriminated against, this is not a collective grievance until others feel the same, and the group is collectively aware of it. Hardin (1995) reasons that repeated social interactions are crucial for developing a shared feeling of discontent. Essentially, by fostering the development of shared grievances in a group population, social interactions help prepare the motivation for violent action. The second and third mechanism of spatial proximity on group mobilization have practical consequences on the feasibility of rebellion. Lichbach (1995) examines the "rebel's dilemma" from a collective action perspective and emphasizes close interaction between participants as an advantage for a rebel organization. This is mainly due to the facilitated flow of information and the exertion of social pressure. The flow of information in a group helps to reduce organizational costs for collective action, which facilitates the organization and coordination of activities. Also, information about the current state of participation in a movement encourages others to join as well (Granovetter 1978; Kuran 1989). At the same time, closely linked participants can monitor contributors and non-contributors much better and thus provide rewards or apply punishments effectively. In general, spatial proximity between participants is likely to have practical advantages for a movement.

In sum, direct social interaction is important for conflict mobilization, and this is also what the empirical analysis in Chapter 4 suggests. However, how does the geographic settlement pattern of a group affect its interaction opportunities, and ultimately the difficulty of mobilization? In the present chapter, I examine the spread of mobilization through the group population on a fine scale, i.e. between small basic geographical units, corresponding to cities or villages. *Within* these small units, I assume that there is a sufficiently close interaction between people

that facilitates mutual radicalization of individuals, but also the monitoring of defectors not willing to join a movement. This essentially means that we can assume that a population in a single unit acts in unison, such that these locations constitute the set of basic units across which mobilization spreads. However, how does mobilization spread *between* these units?

There are different candidate mechanisms that could account for this spread, and we can roughly distinguish between two broad categories. First, non-spatial explanations claim that communication between group members provides sufficient interaction to establish loyalties and facilitate group cohesion (Warren 2007). For these explanations, what matters is the flow of information through media and telecommunication networks, and to a much lesser degree the personal interaction of individuals. If a non-spatial mobilization process was at work, we would see conflict mobilization to spread through the population, but there would be no spatial proximity of the affected locations. This is different from what is typically called a spatial diffusion process (Hägerstrand 1967). Here, the spread through a group population explicitly follows a spatial pattern such that the diffusion typically occurs between spatially proximate locations. In a spatial diffusion process, group mobilization is driven by direct social interaction between individuals rather than indirect communication between group members.

I argue that it is more accurate to conceive of the spread of mobilization as spatial diffusion mechanism. The mechanisms introduced above and in Chapter 2 suggest that spatial proximity of individuals matters for mobilization to spread. This is consistent with evidence on other group processes, as for example the spatial diffusion of innovation. In his famous study, Hägerstrand (1967) showed that the spread of a technological innovation follows a geographic pattern, with the presence of adopters at a particular location making it likely to trigger adoption nearby. If group mobilization is indeed driven by direct social interactions, its diffusion should resemble a spatial mobilization process, with the spread of mobilization from one location to the other driven by the exchange of people between two locations. These people bring with them not only the revolutionary ideas, but also information about current and future rebellious activities. At the same time, they are able to exert social pressure on their peers as to make them participate. In order to approximate the exchange of people between two locations, I draw on Zipf's (1946) results on the intercity movement of persons. He showed that the number of people traveling between two cities is proportional to the populations of both cities, and inversely proportional to the distance between them. Correspondingly, the strength of the social influence between two locations should be greater the more populous these locations are, but decline with distance. This reflects our intuition: If there are two major cities located close to each other, there will be a lot of traveling between them, and revolutionary activity will spread quickly from one place to the other. However, consider the example of two small villages on opposite sides

of a group territory. If a group of people from one village decide to take violent action on behalf of the group, this activity is unlikely to spread to the other village.

Even though it is applied here at a fairly high resolution, the gravity model of interaction has been found to be consistent with a variety of phenomena in the social sciences also at a larger scale. Most of this research has focused on states as the main actors. For example, trade flows between states tend to follow a a relationship given by the gravity model, such that more proximate states trade more (Isard 1954). However, negative interactions between states also appear to be governed by a gravity-like effect. In Chapter 2, I outlined different works that link geographic proximity to a higher likelihood of conflict (Starr and Most 1978; Starr and Thomas 2002; Gleditsch 2002; Buhaug and Gleditsch 2006). Again, the core of the explanation is that geographic proximity facilitates interaction and thus creates opportunities for states to fight. Also, the likelihood of conflict is influenced by the size of states in a dyad. This argument is spelled out in detail in Boulding (1962), where he states that stronger states might be able to project their power over larger distances, thus allowing them to fight wars with more distant states, which again increases the number of potential opponents. A joint test of the distance and size of countries was done by Hegre (2008) who found a strong effect on conflict risk, following a gravity model.

Coming back to the more local interaction I focus on in this chapter, we can hypothesize that mobilization spreads geographically from a few extremist locations through the population of the group, similar to Hägerstrand's (1967) model of the geographic spread of innovation. I implement the influence between these locations as following a gravity model. The next section introduces this model in detail.

5.3 Simulating Group Mobilization in Space

5.3.1 Introducing the Model

The mobilization model is based on the assumption that mobilization geographically spreads through the settlements of a group population. It operates on a set of locations $C = \{c_1, ..., c_n\}$ with given populations $s(c_1), ..., s(c_n)$. The settlements are located at a certain distance from each other, denoted by $d(c_i, c_j)$. At a given point in time, each of these settlements is either in state "mobilized" or "unmobilized". I refer to the set of mobilized and unmobilized cities as M and U, respectively. Obviously, $M \cap U = \emptyset$ and $C = M \cup U$. The model starts with an initial set M_0 of mobilized settlements, and then simulates the spread of mobilization to the remaining ones U. In the following, I introduce the initialization phase and the subsequent mobilization phases in detail.

Initialization phase

During the initial phase of the model, a certain proportion of locations is selected at random to represent mobilized locations. All other locations are set to "unmobilized". The percentage of mobilized locations is kept low: If chosen too high, almost the entire group would be mobilized initially, and no geographic spread would take place. On the other hand, if the initial value is too low, the result would be very few initially mobilized locations, and there would not be any geographic spread, either. I experimented with three different values of this model parameter: 0.03, 0.05 and 0.1. Since the initially mobilized locations are chosen at random, it necessary to perform multiple runs with different random seeds in order to remove the dependence of model results from initial conditions. See below for a description of the model evaluation.

Mobilization steps

In each time step, an unmobilized city is selected randomly from U and is evaluated according to whether it should switch its status to "mobilized". This decision is made based on the influence of the other settlements of the group. As I argued above, the influence of another settlements is proportional to the movement of people. As described in Zipf (1946), the number of people traveling between two cities is proportional to the sizes of the cities, and inversely proportional to the distance between them. In the notation introduced above, the mobilizing influence of city c_i on city c_j is therefore given as

$$I^+_{i,j} = \frac{s(c_i)s(c_j)}{d(c_i, c_j)}. \tag{5.1}$$

However, apart from transmitting radical ideas from one location to another, intercity movement can also inhibit mobilization. In other words, the connectedness of a settlement to many others which do not share rebellious ideas might prevent this settlement from becoming radicalized itself. This inhibiting influence is also hypothesized to be proportional to the exchange of people between the cities as introduced above.

To summarize, the total mobilizing influence I^+ on an unmobilized settlement c is simply the sum of all mobilizing influences

$$I^+ = \sum_{m \in M} \frac{s(c)s(m)}{d(c, m)} \tag{5.2}$$

and correspondingly for the total inhibiting influence,

$$I^- = \sum_{u \in U-c} \frac{s(c)s(u)}{d(c, u)}. \tag{5.3}$$

The probability of the selected settlement c flipping from "unmobilized" to "mobilized" takes into account both the mobilizing and inhibiting influences of all the other group settlements. It is assumed to be proportional to the ratio of the total mobilizing influence I^+ to the total influence $I^+ + I^-$. This mobilization procedure is repeated for a fixed number of time steps, which is a multiple of the total number of cells a group occupies. This ensures that on average, cells of larger groups have the same probability of being activated during the mobilization process as cells of smaller groups. This multiple constitutes the second parameter of the mobilization model. In my experiments, I used three different values (3, 5 and 8), meaning that each group cell would be activated on average 3, 5 or 8 times.

Model Output

The purpose of the model is to examine how well mobilization can spread through the group population. An indicator for this is the proportion p_{mob} of mobilized locations, simply defined as

$$p_{mob} = \frac{|M|}{|C|}. \tag{5.4}$$

In summary, the model works as follows. The behavior is guided by three parameters: First, the proportion of initially mobilized locations (0.03, 0.05 and 0.1); second, the number of times a location is activated (3, 5 and 8); and third, the seed of the random number generator used for this model run. I perform a full evaluation of the entire parameter space, repeating each run for 100 different values of the random seed. This results in 900 realizations of the diffusion process. For evaluation purposes, I computed the average of p_{mob} across all different random seeds. As the tight linking of the model to geographic data (see next section) entails high computational requirements, it was beneficial to keep the number of model runs within a feasible range.[1]

5.3.2 Running the Model on Geographic Data

Computational models have been used for a variety of purposes. Surprisingly, there is little consensus in the community of modelers as to whether a model should be empirically validated, and if so, how this should be done (Moss 2008; Windrum, Fagiolo and Moneta 2007). However, how can models be helpful to social scientists – whose aim is to explain some social phenomenon – if no attempt is made to link the model to reality? To be sure, the linking of a model to the social reality can be done in a variety of ways, and in different degrees of detail. An important distinction can be made according to whether only the model input, the

[1] The full parameter sweep took about 10 days to complete on an Intel Core 2 Duo Processor (2.4 GHz).

model output, or both are used in comparing the model to the real world. An example of the first kind are agent-based models that use interview data from stakeholders to specify decision rules which guide agent behavior (Bousquet et al. 1999). Cederman (2003) is an example of the second kind, since the output generated by his model is compared to an observed empirical regularity: the distribution of war sizes.

The model presented in this chapter is an example of the third kind, since both the model input and the model output are tightly linked to the social reality. Real-world geographic settlement patterns of ethnic groups are used as model input and provide the space in which the group mobilization is simulated. Two types of data are required as model input: data on the primary settlement regions of the marginalized ethnic groups, and data on the population distribution in each group territory. Data on the settlement regions of groups is taken from the GREG project, introduced in Chapter 3. I employed a rasterization procedure to obtain group information for raster cells as fixed spatial units. The result of this process is a raster dataset with a resolution of roughly 30x30 km where the raster cells are occupied by the predominant groups. These raster cells constitute the basic spatial units C on which the model is run. Data on population in the group cells is taken from the *Gridded Population of the World* dataset, released by CIESIN (2005). The distance $d(c_i, c_j)$ between two settlements is computed "as the crow flies" on a spherical earth. Figure 5.3 illustrates the group geography that is used as model input.

5.4 Empirical Test of the Model

Above, I set out the dual validation strategy for the mobilization model, so that both input and output of the model are tightly linked to reality. The initialization of the model with the "real" group settlement patterns constitutes the first part of this strategy, with model input that is directly based on empirical data. This section describes the second part of the empirical validation. Having simulated a hypothetical mobilization process with a computational model, we now must test whether higher mobilization rates as predicted by the model are indeed related to a higher likelihood of conflict. Obviously, the relationship between mobilization rate and conflict is not deterministic, so we do not expect all mobilized groups to exhibit conflict in order for the model to find empirical support. Instead, the relationship should be probabilistic such that a higher group mobilization as computed by the model is related to a higher risk of conflict. I test this relationship statistically, again using the data from Buhaug, Cederman and Rød (2008) on conflict involvement at the group level. This dataset was already introduced in Chapter 4. It codes political status (included vs. excluded), conflict involvement and a set of demographic and geographic variables for a sample of politically marginalized ethnic groups,

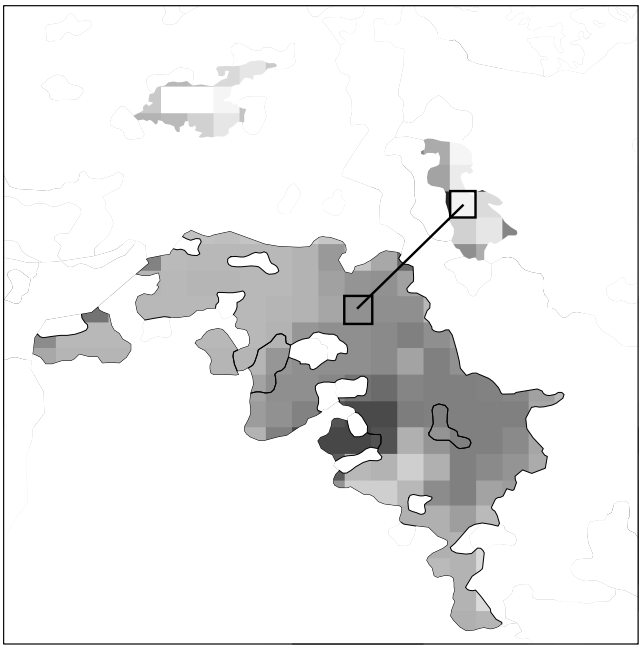


Figure 5.3: Example: Rasterized representation of the Kurds in Iraq. The cells constitute the set of basic units C of the mobilization model. The cell shading indicates the population $s(c_i)$ according to the *Gridded Population of the World* dataset, where darker shading corresponds to a higher number of people. The line shown in the figure depicts the distance $d(c_i, c_j)$ between two cells, measured between their centroids.

currently covering Europe and Asia. In this analysis, I use the cross-sectional version of the Buhaug et al. dataset, where each observation corresponds to an ethnic group.

5.4.1 A Cross-validation Procedure for Parameter Selection

The model as introduced above has two main parameters: the initial proportion of mobilized locations, and the number of times a cell is activated during the model run. Before we proceed to a statistical evaluation of the computed results, we need to select a particular parameter setting to be used for the evaluation. How should this be done? Obviously, selecting the parameters so that they maximize the expected correlation between the mobilization rate and conflict is not particularly informative, as it reduces the whole task to a model-fitting exercise that cannot tell us anything about whether the proposed mechanism has some generality.[2] A better strategy is cross-validation, where the sample is split into n folds. For each of these folds, the model is then run on $n - 1$ folds, and evaluated on the n-th fold (see e.g. Witten and Frank 2005). This

[2]Despite this obvious drawback, there are still examples of this strategy (see e.g. Lim, Metzler and Bar-Yam 2007).

way, we obtain an out-of-sample evaluation of the model; in other words, the set of cases for fitting and testing the model are disjoint.

In order to apply a cross-validation strategy to select the parameters for the evaluation of the mobilization model, I proceed as follows. I randomly split the full sample into two folds. These sub-samples are stratified; in other words, each of the folds contains the same proportion of conflict cases (about 9.5 percent). I then select the parameter values that maximize the correlation between the mobilization rate and conflict on the *first* sample, and use them to compute the mobilization rate on the *second* sample. This procedure is repeated in reverse order, with parameters selected on the second sample used for computing the mobilization rates on the first sample. This procedure ensures that the model parameters are optimized based on a set of cases different from the ones used for evaluation. Having computed the mobilization rate for each group in this way, I test the explanatory impact using regression analysis.

5.4.2 Results

Due to the GREG rasterization procedure applied to the vector dataset, the number of cases in the present analysis is lower than the original one, since very small group occurrences in a country are not included. The final dataset consists of 544 groups from 57 countries, as compared to 867 groups in the original dataset. As mentioned above, I employ a cross-sectional design, comparing groups across the entire post-WWII period. More precisely, the dependent variable takes a value of 1 if a group was in involved in violent ethnic conflict in the post-World War II period through 2005.

The mobilization rates computed by the model range from 0 to 0.82, with a mean of 0.39. In other words, on average about 40 percent of the group cells are mobilized once the model terminates. I first conduct a bivariate comparison, examining how mobilization rates obtained with the model relate to the occurrence of conflict. In order to do so, I split the sample into low and high values of mobilization (median split at 0.56) and count the conflict/no-conflict cases in the respective subsets. Table 5.1 shows the results.

	Low mobilization	High mobilization
No conflict	259	232
Conflict	14	39

Table 5.1: Bivariate comparison between mobilization rates and conflict.

The results in Table 5.1 provide some preliminary support for the proposed mobilization mechanism, as high mobilization rates tend to correspond to a higher conflict risk. Among the 53 conflict cases, more than 70 percent are groups with a high mobilization rate. Whereas

in the subset of groups with low mobilization only about 5 percent are eventually involved in violent conflict, this number almost triples in the high mobilization subset. It is worth taking a closer look at a few cases which the model identifies as low- or high risk cases. Among the correct positive classifications (high mobilization and conflict) are the Ossetians in Georgia. During and after the breakup of the Soviet Union at the end of 1980s, antagonism between the Ossetians and the titular majority, the Georgians, grew increasingly violent up to the level of ethnic civil war (Cornell 2002). The settlement pattern of the Ossetians in Georgia favors quick mobilization: the model computes an average score of 0.78 with a standard deviation of 0.288 across all 100 runs. Also a correctly predicted case are the Kurds in Turkey, who have been in conflict with the Turkish government for many decades. They also receive a high mobilization score of 0.67 with a standard deviation of 0.062. There are also a high number of correct negative classifications (upper left cell in Table 5.1). For example, the Polish community in Lithuania gets a low mobilization rate (0.37, standard deviation 0.35). Even though the Poles in Lithuania have a regional base in the Vilnius region, there are also many settlements close to the border with Belarus (Fearon and Laitin 2006), which accounts for the slow mobilization rate computed by the model. Also among the correct negative classifications is the Uzbek population in Kazakhstan with a mobilization score of 0.49 (standard deviation 0.31). The Uzbeks have a settlement pattern unfavorable for mobilization, as they are spread out along the border with Uzbekistan.

However, Table 5.1 clearly shows that the model gets a number of cases wrong. Among the false negatives is the case of the Chechens in Russia who end up in the "low mobilization" category, but who have been involved in a devastating civil war for years. Still, the Chechens have a mobilization score of 0.52 (standard deviation 0.29), so they just fall short of the threshold for high mobilization. Another case that also fails to be correctly identified are the Tamils in Sri Lanka who are involved in a still unresolved military conflict against the state. Their mobilization score is 0.55 (standard deviation 0.28), indicating again that the model is not far off. There are also a high number of false positives, including for example the Kazakhs, the largest ethnic minority in Mongolia. The Kazakhs appear to have a settlement pattern favorable for mobilization, as the computed mobilization rate of 0.66 (standard deviation 0.168) shows. However, relations between the Kazakhs and the Mongolian government have not shown a trace of conflict (Levinson 1998). Another false positive is the Azeri population in Georgia (mobilization rate 0.59, standard deviation 0.32), who have remained peaceful despite the number of violent conflicts the Georgian government has fought against other minorities (Cornell 2002).

In the following, I will subject the model results to a more thorough multivariate test. Besides the main independent variable of interest, the model also controls for the total number

of group cells, in order to exclude any effects that might be due to the group size rather than mobilization spread, as hypothesized above. Also, I include the following independent variables from the original study: the dyadic *power balance* between the group and government; a population-weighted logged *distance* measure of the group's territory to the capital; and a measure of *roughness of terrain* of the group's territory, computed from elevation data. Furthermore, I add two measures at the country level: logged GDP per capita values and a dummy variable for democracy based on Polity IV, either for 1950 or the year of independence if the state was formed after 1950. I employ logit models to test the influence of mobilization rate on ethnic conflict. Standard errors are clustered by country to control for possible dependency of conflict observations. The results are given in Table 5.2.

	(1)	(2)	(3)
Mobilization rate		2.55**	1.83*
		(0.77)	(0.89)
Total number of cells		-0.00	-0.23
		(0.10)	(0.14)
Power balance	0.72**		0.73**
	(0.12)		(0.15)
Capital distance	0.68**		0.74**
	(0.21)		(0.25)
Mountains	1.24*		1.07
	(0.55)		(0.56)
GPD p.c.	-0.20		-0.18
	(0.20)		(0.20)
Democracy	0.92*		0.97*
	(0.39)		(0.40)
Constant	-2.68	-3.44**	-3.41
	(2.10)	(0.41)	(2.24)
Observations	535	544	535
Log-likelihood	-121.70	-155.15	-113.17

Table 5.2: Logit regression results. Dependent variable: conflict occurrence. Standard errors are given in parentheses. $^{**}p < 0.01$, $^{*}p < 0.05$.

Model 1 repeats the original analysis in Buhaug, Cederman and Rød (2008) with my restricted sample. The results are consistent with those reported in the original article. Power balance has a positive effect, so the larger the marginalized group in relation to the ethnic group in power, the higher the risk of dyadic conflict. The distance of a group from the capital also has a positive effect, as in the original study. Terrain roughness also seems to foster conflict, as indicated by the positive coefficient of the "Mountains" variable. Economic performance as measured by GDP seems to have no effect, but the positive and significant sign of democracy

is somewhat counter-intuitive to our expectations. Model 2 includes the mobilization rate from the computational model, as well as the number of cells as the only control variables. Here, the mobilization rate of a group is positively related to conflict risk and strongly significant, which lends preliminary support to my hypothesized mechanism of group mobilization. In Model 3, I include my variables and the ones from the original analysis. It shows that the effect of mobilization found in Model 2 persists in the presence of the other control variables. The effect of the mobilization rate on conflict corresponds to a three-fold increase in the predicted probability of conflict (from 0.05 to 0.16), when increasing the mobilization rate from the 5th to the 95th percentile while keeping all other variables at their mean values.

5.5 Chapter Summary

In this chapter I have provided a more detailed account of the opportunity mechanism linking geography and conflict, which was introduced in the previous chapter. I proposed a mechanism of ethnic mobilization on a spatial dimension. This mechanism operates on small raster cells as the elementary spatial units of a group. In the model, mobilization initiated by a few extremists spreads geographically to other, previously unaffected settlements. This simulated spread is dependent on the interaction between the settlements, which is approximated by the number of people traveling between the them. For certain groups, the settlement pattern should favor mobilization and therefore conflict involvement of that group. A full test of the proposed mechanism would require data on outbreaks of violence, both over space and time and for different groups. Instead, I used a computational model to artificially "mobilize" a group and to generate a measure for the observed rate of mobilization. The model operates on real geographic data about a group's region and its population distribution. The relation of mobilization speed to conflict was then assessed in a statistical analysis. The results presented here again show that the opportunity-driven mechanism linking group geography and conflict seems to matter a lot. The output from the computational mobilization model is significantly related to conflict risk. More specifically, all others being equal, a group geography that facilitates interaction between group members puts the group at a higher risk of conflict.

This chapter has shown how computational modeling can be used for the empirical testing of hypothetical mechanisms. Although my empirical test only looks at the "input" (group geography) and the "output" (conflict) side of a process (Hedström and Swedberg 1998), the black box in between the two is made explicit. However, the model also generates predictions for the intermediate steps of a process, which could be compared to empirical evidence if data is available. In general, the computational approach to the empirical testing of mechanisms has

some advantages. For example, if one were to revise the hypothesized mechanism, it could easily be implemented and subsequently subjected to empirical testing without a new collection of data. Therefore, a computational model can be considered as an operational specification of a mechanism because it is able to generate the consequences of the mechanism accurately and systematically.

Chapter 6

Settlement Patterns and the Location of Violence

THE preceding two chapters presented an analysis of the impact of settlement patterns on ethnic conflict at the group level. More precisely, by comparing groups' spatial distribution, I showed that the interaction opportunities for group members enabled by geography make certain groups more prone to violence. In this chapter, I narrow the geographic perspective. Rather than studying the conflict involvement of ethnic groups, I examine how settlement patterns determine the location of violence. In other words, this chapter analyzes *where* violence is applied once a conflict has started.

Analyzing the location of violence in civil wars is not new, but has witnessed a surge of interest recently. An example of an earlier analysis of the location of fighting is McColl (1969) who studies the spread of rebellious activity from single rebel strongholds to larger areas. However, due to the development of new approaches, methods and data, new works in the field allow for new insights into the dynamics behind civil war violence. Most importantly, the application of GIS software (see Chapter 3) makes it possible to perform large-N comparisons where studies would normally be limited to a few cases. Buhaug and Gates (2002) is one of the first studies to apply this method, and tries to explain the location and geographic scope of a conflict. More recently, Buhaug and Rød (2006) study the location of civil war violence in Africa by using artificial 100x100 km grid cells as the unit of analysis. At an even finer scale, the development of geographically referenced conflict event data (Raleigh and Hegre 2005) makes it possible to study the determinants of violence down to the level of individual settlements. Using these new data, Hegre and Raleigh (2006) find that violence is likely to occur in regions with an intermediate population density.

However, to my knowledge there have been few systematic attempts thus far to study the

micro-mechanisms between geographic group configurations and violence in ethnic conflicts.[1] An exception is the analysis by Melander (2007) who tries to explain the location of ethnic cleansing. The almost absolute lack of literature on this topic is surprising, given that there is both a theoretical and practical need for research. Theoretically, from the above-mentioned studies, we know how population, local economic performance and other local determinants affect the risk of violence, but little attention has been given to the spatial distribution of groups as the driving force of where conflict erupts. If the geographic pattern of violence is partly affected by group geography – as this chapter argues – studying the location of ethnic violence is key to understanding the strategic motives of actors in these conflicts. Like states as the main actors in international conflicts, ethnic groups might follow strategic considerations in deciding where to fight. These theoretical insights are also relevant for practical reasons. First of all, intervention strategies – whether military or not – would benefit from geographic risk assessments. If certain locations are considered at risk due to their strategic position on the ethnic map, measures could be taken to prevent or mitigate conflict at these locations. No less important is the question of territorial "solutions" for ethnic conflict. For example, Kaufmann (1996) advocates territorial partition as a possible solution to intractable ethnic hostilities. This recommendation has generally been met with criticism. For example, in his statistical test of the partition argument, Sambanis (2000) finds no evidence that partitions help prevent civil wars from recurring. Along similar lines, Laitin (2004) warns against underestimating post-separation conflicts, which are more likely to occur if territorial concentration has been consolidated by the forceful partition of groups. Most importantly, without a detailed understanding of the relationship between the territorial distribution of groups and the application of violence, it is difficult to predict whether externally imposed partitions will lead to a mitigation or even termination of hostilities.

This chapter starts by introducing the theoretical foundations that drive the location of ethnic violence. I argue that competing territorial aspirations of groups determine where violence is applied. The general logic is that violence will occur if a particular location is claimed by more than one group as part of their greater territory. The clashing claims that ethnic groups make over a location is what I call "ethnic contestation" at that location, and it constitutes my key explanation for the effect of group geography on violence. Ethnic contestation makes "regular" war between the groups' armed forces more likely, since groups want to secure control of these units. At the same time, contested units should be the ones that suffer from systematic violence against civilians. More precisely, since these units are of high importance to the competing groups as part of their future territory, this is where ethnic cleansing is likely to occur. The subsequent parts of this chapter present an empirical test of this theoretical argument. An

[1] I do not count here the analysis by Lim, Metzler and Bar-Yam (2007) since case selection and method are problematic, which makes the results questionable.

analysis of the location of ethnic violence requires much more detailed data of ethnic settlement patterns than the group-level analysis presented in the previous chapters. For that reason, I test my hypotheses on a sample that is limited in geographic scope, but measures group geography and violence at a fine-grained level. My dataset on settlement patterns in Bosnia, introduced in Chapter 3, serves as the basis for this study. Based on these data, I compute a measure of the strategic aims of groups, which is then tested in a spatial regression analysis of the occurrence of ethnic violence at the municipality level. The analysis of the consequences of violence with respect to the spatial distribution of groups is continued in the next chapter, where I examine the effect of conflict on territorial changes.

6.1 Clashing Territorial Claims and the Application of Violence

How does territory affect the application of violence? In Chapter 4, I argued that there are two potential mechanisms linking the spatial distribution of a group to its potential for conflict – first, a motivation mechanism where territorial concentration makes groups more likely to fight for it, and second, an opportunity mechanism that relates concentration to the group's potential for mobilization. The two previous chapters have shown that there is empirical support for the latter. In the present chapter, however, I move from a monadic setting of groups as the unit of analysis to a dyadic one, studying how territorial relationships between groups affect conflict. From this dyadic perspective, it is worth going back to the motivational arguments about the importance of territory to groups. Chapter 2 mentioned the different material- and non-material reasons that determine the close link between groups and "their" land. Territory matters for groups because the control of it secures access to resources (Hensel 2000). Moreover, a group's settlement area is often an integral part of its identity (Smith 1986). The control of a "homeland", as Toft (2003) emphasizes, is crucial for the survival of a group identity.

Under which circumstances can territorial issues incite violence? Chapter 2 introduced two types of ethnic conflict, which I labeled "asymmetric" and "symmetric". For each type of conflict, the role of territory in determining violence is likely to be different. In an asymmetric conflict that corresponds to the classic definition of a civil war, a comparably small and weak ethnic minority group is pitted against the government of the state it lives in. The government exerts authority over the group's territory, which is challenged by the group. In this type of ethnic conflict, it is not the *extent* of the "homeland" of the group that is under dispute, but the *control* over it. Consider for example the ethnic conflicts in Georgia since the end of the Cold War. Both the Abkhaz and South Ossetian minorities live in "their" administrative units, the

borders of which are reasonably well defined. The conflict between these minorities and the Georgian government is mainly about the control of these units (Cornell 2002).

However, this is different for "symmetric" ethnic wars, which resemble an all against all situation with equal actors. During periods of state failure when there is no central authority guaranteeing the peaceful co-existence of groups, the groups themselves are responsible for their own security (Posen 1993*a*). Other groups are perceived as a threat, and preemptive measures are taken as a reaction. In such a setting, the territorial logic is different from the aforementioned asymmetric type of conflict. In the absence of a strong state that guarantees inter-ethnic boundaries, the definition, securing and defense of these boundaries is likely to become an issue of high priority for groups. This is especially problematic in cases where group territories are not clearly demarcated: in these cases, groups will fight over the *extent* of territory (i.e. its boundaries), rather than *control* over a predefined area.

The process of the violent formation of inter-ethnic boundaries in symmetric ethnic conflicts is the focus of this chapter. More precisely, in a security dilemma condition of all against all, I examine which locations ethnic groups choose to fight for. According to Posen (1993*a*), security dilemma conflicts occur during the breakup of multi-ethnic states, and groups fight with the prospect of eventually creating new state entities from their territories. During this process, the question is what the spatial extent of these entities will be. Key to the understanding of ethnic boundary formation is an assumption that might seem tautological at first: Ethnic groups recognize "their" territory by the location of their ethnic kin. Consequently, groups will try to align a territory that includes all locations in which members of their group live. This strategy can be easily implemented as long as we deal with cases in which ethnic regions are clearly separated, but is likely to prove difficult in ethnically mixed areas. What if an area is inhabited by more than one group? Since more than one group will claim the territory as theirs, disputes over the ownership of the territory are likely to arise.

The general question this chapter examines is how these clashing territorial claims relate to the occurrence of violence. In a security dilemma situation, groups fight with the goal of creating ethnically pure successor states. However, for these states to be viable, their boundaries must follow some basic principles, most importantly that the territory be contiguous. It is obvious that a state with territorially isolated enclaves will face enormous difficulties in defending them (Posen 1993*a*; Kaufmann 1996). Also, a contiguous territory ensures economic benefits because of uninterrupted communication and transportation routes (Melander 2007). For that reason, groups will take into account the position of a location relative to the larger group territory. Consider a province with a high population share of a particular group A. If this province is situated close to a larger region where group A is clearly dominant, A would be more likely

to claim that province than if the province were surrounded by mixed areas and thus could not easily be added to A's larger territory. Consequently, when choosing which of its populations to include in the future state, an ethnic group should tend to favor those that constitute a core territory, and those that are close to it. This strategy will not lead to disputes if only one group raises claims to a particular location. However, if there are two (or more) groups demanding control over a unit, this might trigger the process of violent conquest and cleansing described above. What I call *ethnic contestation* at a location is the degree to which two or more groups aspire to occupy the same territory.

How does ethnic contestation trigger violence? For each of the groups laying claim to a particular location, the goal is to eventually add that location to their greater territory. To this end, two steps are necessary. First, one group must establish control over the unit, and second, the winning group must cleanse the unit of the enemy group population. At each of these stages we are likely to see a different type of violence. The first stage is characterized by violence between armed groups. If groups disagree with regard to the ownership of a unit, both sides are likely to bring in their military forces to take over the unit. The first stage of these violent confrontations over territory therefore resembles interstate conflicts, with armed forces trying to secure territorial control (Holsti 1991; Huth 2000). Consequently, in locations where group demands clash, we should see two-sided violence between armed forces.

However, when control of a location has been established, the process of ethnic cleansing begins (Bell-Fialkoff 1993). Groups will try to set the boundaries of their territory such that they include all of their kin, but at the same time as low an alien population as possible. Ethnic cleansing strategies are employed to achieve the latter. The cleansing of populations from an area can take a variety of forms, from active discrimination of single persons – short of physical violence – to the genocidal extermination of entire populations (Mann 2005, p.12). Indeed, violence plays a crucial role in ethnic cleansing. In contrast to two-sided violence which is likely to occur during the conquest of a location, the violence at this stage is one-sided, perpetrated by the armed forces of the dominant group against the civilian population of the group to be expelled.

We can now state the relationship between ethnic contestation and the two types of violence described above. First, ethnic contestation should lead to two-sided violence, since the groups that consider a location to be valuable for them will try to secure control over it. As outlined, this should lead to violent interactions between the groups' armed forces. Therefore,

> H1: The occurrence of two-sided violence should be higher at locations with a high degree of ethnic contestation.

A similar relationship should hold for one-sided violence. If locations are likely to be added

to a larger territory of a group, the dominant group will try to expel members of enemy groups. As described above, this can happen by different strategies, one of which is violent ethnic cleansing. One-sided violence and violent ethnic cleansing are not exactly the same, as one-sided violence can also be carried out for reasons other than the creation of ethnically homogenous areas. In general, however, we should see more violence against civilians at locations where there is high ethnic contestation. Correspondingly,

> H2: The higher the degree of ethnic contestation at a location, the more one-sided violence we should see.

According to the theoretical perspective presented in this section, territorial configurations of ethnic groups where violence is applied. In short, this process works as follows: In a security dilemma situation, ethnic groups aim to create ethnically pure territories that include all their ethnic kin. This causes territorial ethnic contestation if two groups claim a location to be part of their territory. Ethnic contestation leads to violent military confrontations between groups in their attempts to secure control over the unit (H1). Similarly, ethnic contestation favors violence against civilians as a means to cleanse the unit (H2). I now proceed to the empirical analysis of these considerations.

6.2 Ethnic Contestation and Conflict in Bosnia

This section subjects the hypotheses presented above to an empirical test. I start with a description of the dataset and the indicators, and then proceed to the regression analysis.

6.2.1 Case Selection and Data

For my empirical analysis, I select the conflict in Bosnia from 1992-1995 for two reasons. First, the conflict is a prime example of a territorial ethnic conflict of all against all, with three major groups fighting for territorial gains in the absence of a powerful central authority. The second reason for selecting Bosnia is data availability. In order to examine changes in the ethnic map, my analysis requires both pre- and post-conflict data on the spatial ethnic distribution. As discussed in Chapter 3, many existing datasets on ethnicity are not sufficiently detailed to allow for a fine-grained analysis of settlement patterns and the location of conflict. However, due to the inclusion of ethnic variables in the state census, the case of Bosnia provides us with information on group distributions at the municipality level, which I linked to a map of these municipalities as described in Chapter 3. The high data requirements rule out other potential

cases my analysis could have been applied to. For example, the conflict in the Caucasus region of Nagorno-Karabakh largely displays features of an symmetric ethnic conflict between the Armenian and Azeri populations (Melander 2001). The 1915 Armenian genocide in Turkey is another example (Mann 2005). However, since both of these cases lack precise data about the spatial distribution of ethnic groups, the Bosnian case is the best possible choice.

Estimates for the conflict activity in a particular unit were obtained from ACLED, the *Armed Conflict Locations and Events Dataset* (Raleigh and Hegre 2005). This dataset lists reported confrontations between the fighting parties in a civil war, along with the date and the spatial coordinates of the event. ACLED distinguishes between different types of events. Types 1–3 includes confrontations between armed groups, both with or without territorial transfers. I used events from these categories to compute my variable for two-sided violence. Type 7 in ACLED corresponds to events of one-sided violence, perpetrated by armed forces against non-combatants. The events of this category constitute the basis of my one-sided violence variable. Using GIS software, I counted the number of events for each municipality in order to obtain a measure for one- and two-sided violence. For the former, the number ranges from 0 to 31, with a mean of about 1.1. The latter ranges from 0 to 34 with a mean of 3.7.

6.2.2 Computing an Indicator of Ethnic Contestation

Strategic ethnic contestation arises when two ethnic groups lay claim to a region because they consider it to be part of their greater territory. Thus, the computation of an indicator for ethnic contestation consists of two steps: First, we must measure the strategic importance of a unit for each of the ethnic groups. This number should indicate how well the given unit fits into the larger territory of the group. Second, we must combine the measures of the territory's importance to each of the individual groups into a single indicator for strategic ethnic contestation in a unit. This indicator should detect a clash in group demands, i.e. if more than one group lays strategic claims to this unit.

I first turn to the question of how to measure the strategic importance of a unit to a particular group. A unit has strategic importance to group A if it can easily be added to A's larger group territory. This is only the case if (i) there are other significant populations of A living nearby to which the respective unit can be added, and (ii) if the unit itself has a sufficiently large population of A. By the same token, the unit should have lower importance to A if it is primarily surrounded by populations of some other group B, or the unit itself has a high share of B. I therefore measure the strategic importance of a unit to A by multiplying two factors: first, the average of A's population shares in the unit's neighbors, and second, the population share of A in the unit itself. This measure is computed for each unit and each of the three ethnic groups

in the sample. As expected, the priority levels we obtain for the three groups are negatively correlated. Serb and Croat claims correlate at -0.40, Muslim and Croat claims at -0.38, and Serb and Muslim claims at -0.48. These correlations indicate that on average, the claims of groups do not overlap: if a unit has a high priority for one group, it receives a low priority for the other groups. The scatter plots in Figure 6.1 plot the importance scores of groups against each other. Municipalities that receive high scores for two groups at the same time indicate clashing territorial claims. The plots reveal that the number of contested provinces varies for different group constellations: Whereas there are many units with mutual claims by Serbs and Muslims (upper left panel), this number is much smaller for Croats and Muslims, and Croats and Serbs. These figures suggest, at least initially, that most conflict potential was present between Serbs and Muslims. In general, however, the scatter plots show that are a substantive number of units where there are clashing group claims. I now turn to the calculation of an aggregate indicator measuring this territorial contestation.

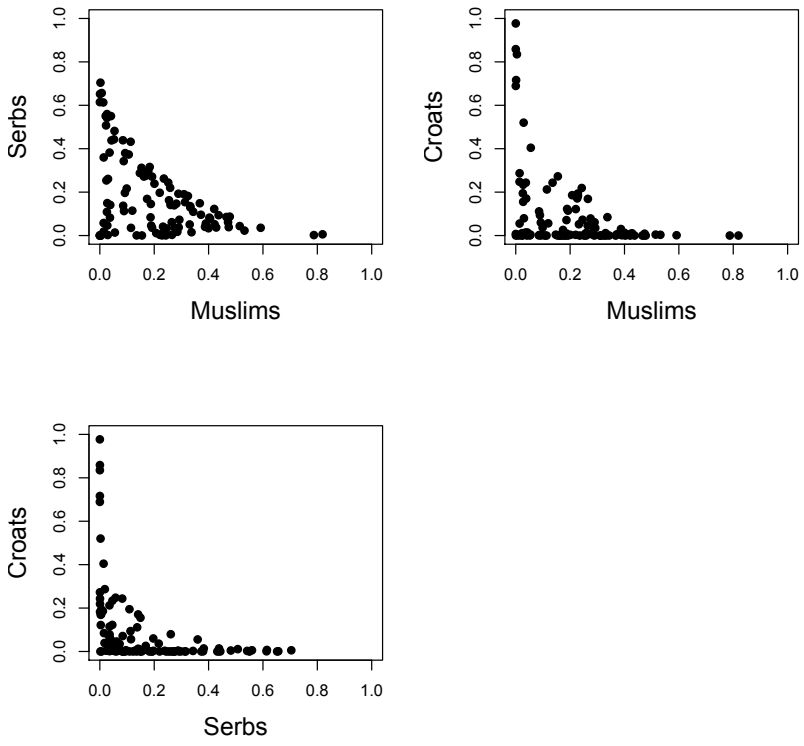


Figure 6.1: Scatter plots of the strategic importance scores of the municipalities in 1991. Dots toward the center of the coordinate system indicate that the unit is considered important by two groups, and is therefore likely to be contested.

The importance scores must now be combined into a single indicator of clashing claims.

A unit should be under dispute if it has a high importance to more than one group. For that reason, I select the two highest importance scores for a unit across all ethnic groups, and use the product of these two scores as my indicator of strategic ethnic contestation. High values of this variable suggest that there are two ethnic groups that consider the respective unit to be part of their territory. The strategic contestation indicator ranges from 0 to 0.06, with a mean of about 0.02.

A short example illustrates the logic behind the strategic contestation indicator. The municipality of Zvornik in eastern Bosnia was one of the first locations to see violence between Serbs and Muslims during the Bosnian war in early April 1992 (Burg and Shoup 1999, p.129). Zvornik has eight neighboring municipalities in Bosnia (see Figure 6.2). Since some of these have high population shares of Serbs (e.g. Bijelina and Sekovici), and Zvornik itself has a Serb share of almost 0.5, the municipality is of high importance to the Serbs (importance score 0.19). At the same time, Zvornik also borders some Muslim-dominated municipalities (e.g. Kalesija and Bratunac) and has a high share of Muslims (slightly more than 0.5). For that reason, it is likely to be also claimed by this ethnic group (importance score 0.29). The multiplication of the Serb and Muslim importance scores for Zvornik results in a strategic contestation value of 0.056 and is among the ten highest scores in the sample.

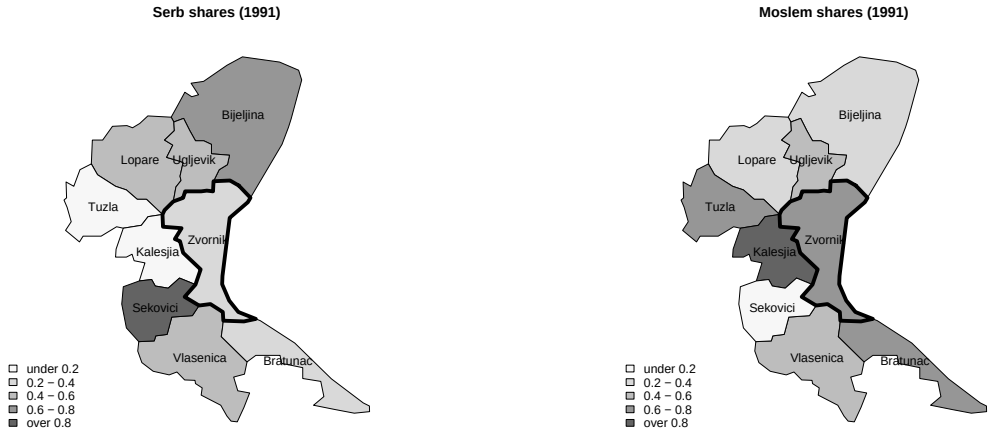


Figure 6.2: Example of the computation of the strategic contestation indicator. The color shading indicates the proportion of the respective group in a municipality, for Serbs (left) and Muslims (right).

6.3 Results

Having introduced the data and indicator for ethnic contestation, I now turn to an empirical test of the relationship between contestation and conflict.

6.3.1 Ethnic Contestation and Two-sided Violence

According to the theoretical discussion presented above, we should see more two-sided violence in units with higher ethnic contestation. For a first visual inspection, Figure 6.3 shows the geographic distribution of the contestation scores, with an overlay of the two-sided conflict events from ACLED. Darker colors correspond to greater ethnic contestation in a municipality. According to H1, we should observe a higher likelihood of two-sided violence in units with darker colors. This expectation seems to be confirmed by Figure 6.3, which shows that most of the units with high contestation (dark) show more conflict activity.

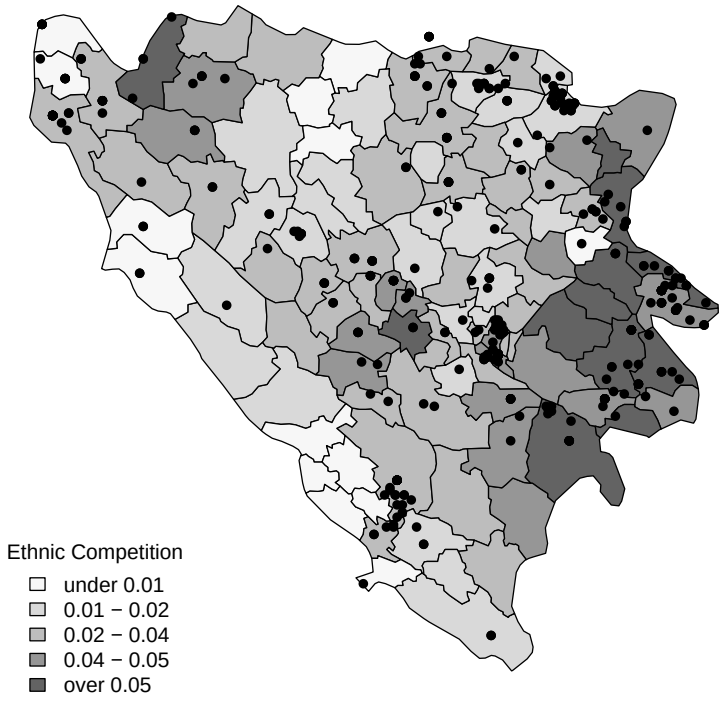


Figure 6.3: Ethnic contestation scores for 1991. Darker colors correspond to higher scores. The dots indicate the location of two-sided conflict events from ACLED.

In order to provide a more thorough test of H1, I employ regression analysis with the number of events in a unit as the dependent variable.[2] This variable is likely to show overdispersion, so

[2] All models estimated using R 2.7.1 and the *Zelig* package (Imai, King and Lau 2006).

a negative binomial model is to be preferred over a Poisson model. I control for the effect of the local ethnic configuration of a municipality in 1991 by including a measure of ethnic diversity computed using the ethno-linguistic fractionalization index ELF (Taylor and Hudson 1972). Note that the criticism on the ELF index mentioned in Chapter 2 applies to the country-level version of this index, which hides much of the intrastate variation behind a single number. In the present application, the ELF is perfectly appropriate, since it measures ethnic diversity at the local level and therefore is able to capture this variation. My expectation is that ethnic diversity should have a positive effect on two-sided violence. I also control for whether a municipality borders regions outside Bosnia, since direct proximity to either Serb or Croat territory might increase the conflict propensity of a municipality. Furthermore, we should expect conflict to be more likely in more populous units, so I include the logged unit population in 1991 as an additional independent variable. A further issue needs to be addressed in the analysis. We must assume that conflict is spatially dependent, i.e. the number of conflict events in a unit is to a certain degree determined by the number of conflict events in its neighboring units. To deal with this dependence, I include a spatial lag in the model (Ward and Gleditsch 2008). The spatial lag in my model is computed as the average conflict count of a unit's direct neighbors. Figure 6.1 reports the results.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	−7.73	1.76	−4.40	0.00
Ethnic contestation	16.06	8.35	1.92	0.05
Local ethnic diversity	2.99	0.96	3.10	0.00
Border unit	0.57	0.28	2.08	0.04
Population (log)	0.60	0.17	3.51	0.00
Spatial lag of DV	0.12	0.05	2.64	0.01

Table 6.1: Negative binomial regression results. Dependent variable: Number of two-sided violence events (N=109).

The regression analysis confirms my expectations. Units with a higher degree of contestation show significantly more conflict activity. When increasing ethnic contestation from the empirical minimum (0.00) to the maximum (0.062), the expected number of two-sided conflict events increases by about 4. However, the occurrence of two-sided violence also depends to a large degree on the local ethnic make-up, with more diverse units experiencing more violence. Municipalities at the border of Bosnia have a higher likelihood of conflict, and the same holds for more populous municipalities. We also see that conflict exhibits a high degree of spatial correlation: The coefficient for the spatial lagged dependent variable is positive and strongly

significant.

As the map in Figure 6.3 shows, the dependent variable in this analysis has a high proportion of zeros, which might bias the results of the negative binomial model. In order to test for this, I also estimate a zero-inflated negative binomial model (results not shown).[3] I do not observe a change in the direction or significance level of any effect, with the exception that the border unit variable is only significant at the 0.1 level.

6.3.2 Ethnic Contestation and One-sided Violence

For a test of H2, I repeat the above analysis with the one-sided violence count as the dependent variable. Again, I start with a visual inspection of the data. Figure 6.4 shows the geographic distribution of strategic ethnic contestation, again with an overlay of the one-sided conflict events in ACLED.

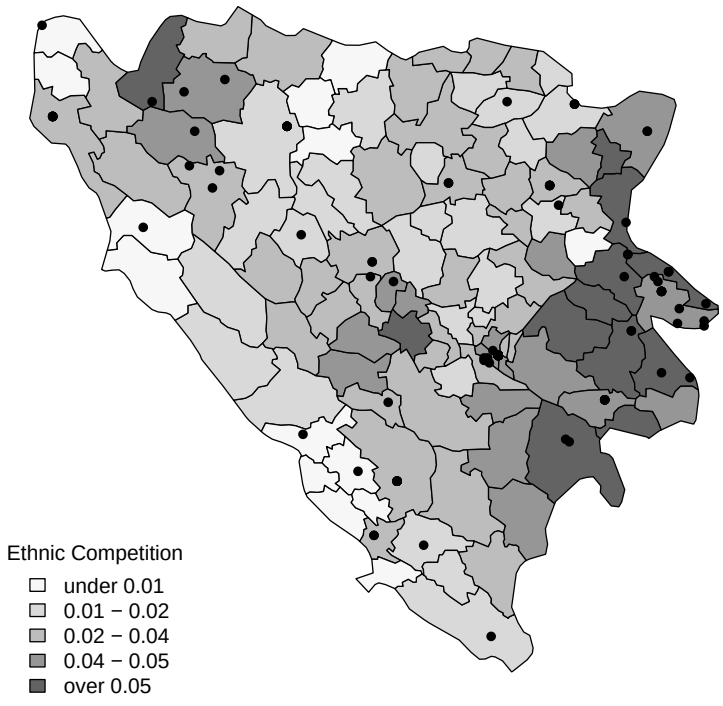


Figure 6.4: Ethnic contestation scores for 1991. Darker colors correspond to higher scores. The dots indicate the location of one-sided conflict events from ACLED.

According to H2, high levels of contestation should be related to increased one-sided violence. Visually, this relationship seems to be supported; we see a lower occurrence of one-sided

[3] Model estimated using R 2.7.1 and the *pscl* package (Zeilis, Kleiber and Jackman 2007).

violence in areas with light shading. I employ regression analysis to provide a more reliable test of this relationship. Again, I use a negative binomial model with the one-sided event count as the dependent variable, and a spatially lagged dependent variable to control for spatial dependence. Table 6.2 presents the results.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	−20.29	3.21	−6.32	0.00
Ethnic contestation	29.06	14.14	2.06	0.04
Local ethnic diversity	−0.69	1.58	−0.44	0.66
Border unit	0.56	0.41	1.37	0.17
Population (log)	1.78	0.31	5.82	0.00
Spatial lag of DV	0.30	0.15	2.08	0.04

Table 6.2: Negative binomial regression results. Dependent variable: Number of one-sided violence events (N=109).

The model shows two major results. First, in line with H2, ethnic contestation is positively and significantly related to one-sided violence, in other words, units with high contestation as captured by my indicator show significantly higher levels of violence against civilians. The effect of ethnic contestation corresponds to an expected increase by roughly 2 conflict events when increasing contestation from the minimum to the maximum, while holding all other covariates at their means. The second finding pertains to the impact of the local measure of ethnic diversity. As I shown above, ethnic diversity leads to a higher occurrence of two-sided violence. In other words, besides the strategic ambitions of groups as captured by my contestation indicator, the local ethnic configuration explains a great deal of variance in two-sided violence. This stands in stark contrast to one-sided violence. Here, local ethnic diversity turns out to have no discernible relationship to the occurrence of one-sided violence, whereas ethnic contestation is a strong predictor. This result is a first indication of the strategic nature of civilian targeting: If a unit is of strategic importance, the group will make an attempt to induce population changes in that unit. Violent military confrontations, however, seem to be driven by strategic aims, but also by the local ethnic diversity of a municipality. Again, I repeat this analysis also with a zero-inflated negative binomial model. It confirms the findings on the above presented model; the only observed difference is that the ethnic contestation variable is now significant only at the 0.05 level.

6.4 Chapter Summary

Settlement patterns affect conflict in a variety of ways. In the two previous chapters, I showed how the spatial distribution of a group affects the risk of a group being involved in conflict. This chapter has shifted the level of analysis to a finer level and analyzes how settlement patterns affect the location of violence. My analysis of the Bosnian civil war suggests that ethnic conflicts follow a logic that relates violence to the territorial aspirations of ethnic groups. Locations with a high level of ethnic contestation tend to see more confrontations between armed forces as groups struggle for control of a unit, but these locations are also more susceptible to one-sided violence against civilians. The mechanisms I claim to be at work are not applicable to all kinds of ethnic conflicts. The strategic application of violence as a means of securing territory is most likely to happen in symmetric ethnic conflicts, where there is an all against all struggle of groups fighting with the prospect of creating new independent states from their territories.

This and the previous two chapters assume a uni-directional perspective in terms of causality: I treat ethnic settlement patterns as a causal precondition of conflict. In Chapters 4 and 5, I show that group geography influences a group's ability to coordinate for collective action. If the settlement pattern favors mobilization, conflict is more likely. This chapter continues along similar lines, but at a finer resolution, and argues that group geography creates strategic incentives for groups to apply violence at particular locations. The next chapter reverses this perspective and examines the effect of conflict on settlement patterns. How does conflict change the ethnic map? This perspective is largely missing in the scholarly literature, although it is intimately related to ethnic conflict.

Chapter 7

The Endogeneity of Group Geography and Conflict

THE study of civil war has generally treated geography as a factor exogenous to conflict. For example, we know that rough terrain (Fearon and Laitin 2003), the occurrence of diamonds and oil (Lujala, Gleditsch and Gilmore 2005; Lujala, Rød and Thieme 2007), and the distance to a border (Buhaug and Rød 2006) affect the likelihood of conflict. Up until now, this dissertation was no exception: In the previous chapters, I showed that the geographic settlement patterns of ethnic groups influence conflict at various levels. For ethnic groups, the spatial distribution of their members determines interaction opportunities and thus facilitates mobilization (Chapters 4 and 5). From a more local perspective, the previous chapter showed that fuzzy territorial boundaries between groups create strategic opportunities for groups to apply violence at particular locations. However, in the preceding chapters I treated the nature of the main actors, the ethnic groups, as constant.

The present chapter will relax this assumption. Why? The short answer is that treating actors as constant hides much of the dynamics that occur during conflict. This observation is not limited to the study of intrastate conflict between groups. For example, theorizing in International Relations frequently departs from the assumption of fixed actors. This literature has traditionally dealt with the study of states as the main actors, and has given much attention to their cooperative and conflictive interactions (see e.g. Waltz 1979). The reliance on a set of largely constant actors allows for parsimonious and elegant theories, but may limit the scope and accuracy of the theoretical framework. However, more and more attention has been given to actor changes in the international system, acknowledging these changes as an inherent property of the system. For example, Gilpin (1981) distinguishes between (i) *relational* changes that affect the interactions between states, (ii) *systemic* changes that are characterized by transfor-

mation of states vis-à-vis others, and (iii) *systems* changes that correspond to fundamental shifts in the nature of actors, as the disappearance of empires.

This chapter examines what Gilpin's *systems* change corresponds to in the context of intergroup interactions. More precisely, I study the transformation of the main actors – the ethnic groups. Following Wendt (1994), Cederman and Daase (2003) suggest a distinction between transformations of the *social* and *corporate* identities of groups. The social identity corresponds to the group's shared beliefs and ideas, whereas the corporate identity "constitutes a group's very existence and its extension in time and space" (Cederman and Daase 2003, p. 7). In fact, corporate identity transformations have been studied in detail by Bremer and Mihalka (1977) and later by Cederman (2002*a*, 1997), who examine the emergence and disappearance of state on a geographic dimension. As a consequence, treating a group's geography as static, as I have in the previous chapters, corresponds to the negligence of corporate identity transformations. However, these transformations are important, as they are particularly likely to occur within the context of violent conflict. Again, it is worth drawing an analogy to international war. The literature on international conflict has repeatedly emphasized the relationship between territory and war. Holsti (1991) shows that territory has been the most frequent issue of interstate war since the Peace of Westphalia. However, as much as territory spurs violence, wars in turn have an impact on the layout of the international system and oftentimes change the geographic shape of states. Tilly (1992) concisely summarizes this as: "War makes states, and vice versa" (p. 67). Being engaged in external wars, he argues, transforms the institutional structure, but also the geographic shape of states. Clearly, conflict is one of the most important driving forces behind corporate identity changes in the international system.

Accordingly, in the context of international conflicts, we can speak of territory and violence as being endogenous: Territorial configurations trigger war, but war also changes territorial configurations. However, is there a similar logic in intrastate conflicts? In this chapter, I examine the endogeneity of territory and conflict in ethnic conflicts. In doing so, I build on and extend the theoretical discussion of the previous chapter, in which I argued that groups fight with the aim of creating contiguous, ethnically pure territories. If groups can only partly achieve this aim, the effect of violence should be a strategic unmixing of the system, such that there are less competing territorial claims of ethnic groups after the conflict. However, the fact that such a segregation process occurs in times of conflict does not yet tell us much about the particular role that violence plays in triggering population changes on the ground. Again, I distinguish between the role of two-sided violence between the groups' armed forces, and one-sided violence against civilians. I argue that the former is used as a means of establishing control over territory. Once this has been achieved, one-sided violence against civilians is applied as a strat-

egy of ethnic cleansing by enforcing migration or exterminating people. The remainder of this chapter presents an analysis of changes in ethnic settlement patterns in Bosnia during the civil war. As was mentioned in Chapter 3, for this to be possible, we require both pre- and post-war population estimates. For Bosnia, these numbers are available at the municipality level and thus allow for an assessment of the impact of violence on population changes.

7.1 Internal Conflict and Territorial Changes

How does violence change the spatial distribution of ethnic groups? As alluded to in Chapter 1, violent conflict around the world continues to be one of the most important determinants of forced migration (Salehyan 2007). However, here we need to distinguish between cases in which migration occurs as a (no less grave) side-effect of conflict, and cases in which the explicit aim of the conflict is the expulsion of a population. The latter category is the focus of this chapter. There are a wide range of cases of violent ethnic cleansing, with episodes varying in scope and strategy. The Nazi campaign of ethnic cleansing in Europe, committed mostly against the Jewish population, was characterized by an immense number of casualties, but also by its thorough organization on the side of the perpetrators (Mann 2005). Many recent ethnic conflicts have also included episodes of significant ethnic cleansing, such as the Rwandan civil war (Gourevitch 1999), the conflict in the Abkhaz region in Georgia (International Crisis Group 2006), and the Serbs' atrocities in Kosovo (U.S. Department of State 1999). However, even though there is a large body of literature examining the causes of these events, much less has been written about the spatial dynamics of population movements during these conflicts. The purpose of this chapter is to fill this gap and to explain whether and how the application of violence induces changes in the spatial distribution of ethnic groups.

As the previous chapter has shown, groups seek to create an ethnically homogenous territory. In mixed regions, however, this is not possible without population adjustments, so the strategy of ethnic groups is to remove the alien group from the territory they consider to be theirs. This can happen via different strategies, from non-violent cultural suppression to forced migration or even the killing of entire populations (Mann 2005, p.12). During ethnic conflict, we typically observe the whole spectrum of forceful cleansing. This can be thought of as a segregation process, where populations separate into more homogenous clusters – similar to what Schelling (1971) describes in this famous model of neighborhood segregation. As outlined in Chapter 5, Schelling imagines the formation of ethnically segregated population clusters to be the result of individuals' preferences regarding their neighbors. During ethnic cleansing, however, although the result is likely to be a segregated environment, the process by which this

separation is implemented is different, as migration is induced by the application of violence. Since local territorial disputes are at the root of conflict as shown in the previous chapter, the effect of these population changes should (at least partly) reduce the conflict susceptibility of the system; in other words, the territorial incentives for groups to apply violence should be lower. Again using my notion of territorial contestation as introduced in the previous chapter, we should expect contestation to be systematically reduced. I therefore hypothesize that

> H1: Conflict leads to a system with lower ethnic contestation.

Note that H1 only considers the short-run territorial incentives for conflict. A decrease in contestation does not mean that conflict susceptibility decreases in the long run. In fact, settlement pattern changes induced by violence may spur future wars, for example due to the expelled group's feelings of hatred and revenge towards their enemy (Petersen 2002). In general, Laitin (2004) advises caution against territorial solutions and points to the future costs they can entail.

The previous hypothesis examines the system-wide effects of conflict on ethnic distribution. However, what explains the local changes towards less territorial contestation on the ground? As argued above, overall ethnic contestation decreases because of a strategic unmixing of group populations. Although violence is not the only way to induce ethnic cleansing, ethnic configuration changes can be implemented by employing violence strategically. As described above, we must distinguish between the effect of two-sided violence as a means of gaining control over a unit, and one-sided violence as a way of cleansing the unit of an alien population. The occurrence of two-sided violence indicates the struggle between groups for control of a unit. Repeated military confrontations between groups therefore mean that control over a unit has not been established, because fighting should stop if one group takes over the unit. Without clearly established control of a unit, we expect that fewer systematic attempts will be made to ethnically cleanse the unit, which results in a smaller decrease in ethnic contestation. I therefore hypothesize that

> H2: The decrease in ethnic contestation should be lower at locations with a high occurrence of two-sided violence.

One-sided violence, however, should have the opposite effect on population changes. We can assume that violent cleansing strategies are among the most effective in creating ethnically "pure" territories. One-sided violence against civilians is the most extreme manifestation of a group's attempt to achieve ethnic homogeneity, so we assume that high occurrences of one-sided violence should be related to more frequent attempts to achieve ethnic homogeneity, and correspondingly, low ethnic contestation. In short,

H3: The decrease in ethnic contestation should be higher at locations with more one-sided violence.

An example illustrates how the application of one-sided and two-sided violence is responsible for inducing population changes. The municipality of Srebrenica in Eastern Bosnia has received sad publicity due to the genocidal events that occurred there in 1995. Srebrenica had a pre-war Muslim population ratio of about two-thirds, and since it is located close to the border with Serbia, was at a particularly high risk of being attacked by Serbs. The city was the location of repeated confrontations between Muslims and Bosnian Serbs already in 1993 (Burg and Shoup 1999). However, due to the presence of UN troops under the UNPROFOR mission, the Serbs did not manage to take over the city until 1995. As a result of the continuous decrease in UN troops over the years, the Serbs were finally able to bring Srebrenica under their control in early July 1995. Once control was established, the Serbs conducted large-scale killings of the Muslim population, which until now constitutes the "largest such incident in Europe since World War II" (Burg and Shoup 1999, p. 325).

By examining the consequences of violence on the spatial distribution of groups, I continue the causal process described in the previous chapter. If the attempts of groups to create ethnically pure territories are only partly effective, conflict should reduces the ethnic contestation in the system (H1). Local population changes toward lower ethnic contestation should occur where one group has taken control. Many military confrontations between groups indicate that no group has taken over, and are therefore related to a smaller decrease in ethnic contestation (H2). Violence against civilians, however, is an indicator of forceful cleansing and should therefore be related to greater decreases in ethnic contestation (H3). The next section subjects the hypotheses to empirical scrutiny.

7.2 Results

7.2.1 Conflict Reduces Ethnic Contestation in the System

As shown in the previous chapter, ethnic contestation increases the likelihood of both one- and two-sided violence. Correspondingly, if conflict in general is successful as a means of resolving territorial issues, we should see decreasing ethnic contestation during the conflict. In order to test for this, I compared pre- and post-conflict ethnic contestation scores. Post-conflict strategic ethnic contestation is computed as described in Chapter 6, but using the 2000 population data. Before looking at the overall distribution of contestation, I again do a pairwise comparison of the importance scores, similar to Figure 6.1. Figure 7.1 shows pairwise plots of the group's

importance scores for the Bosnian municipalities. Post-war (2000) scores are shown as solid black dots, and the grey dots indicate the pre-war scores for easier comparison. The plot shows that the number of contested units (those closer to the center of the coordinate system) has decreased significantly.

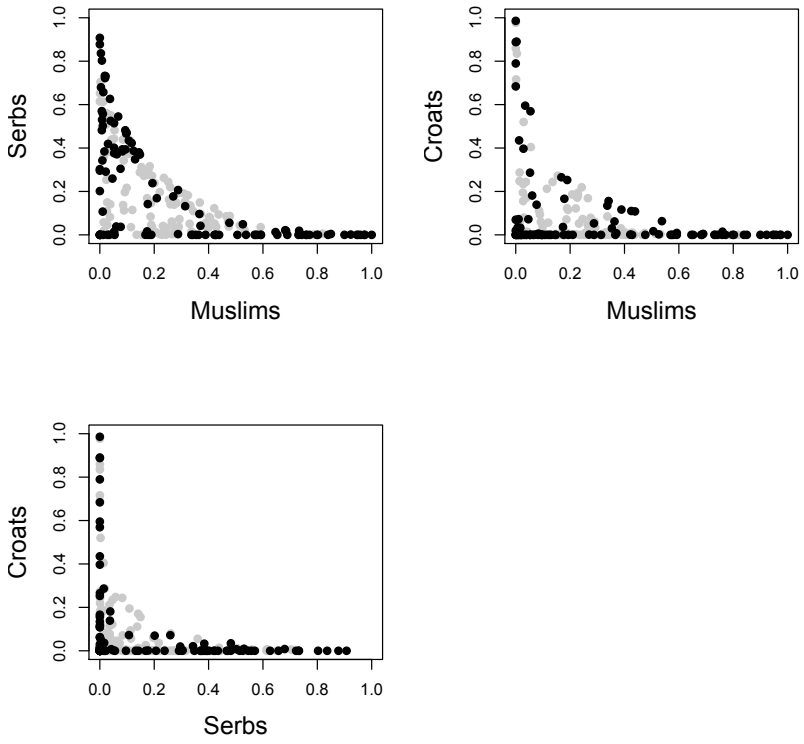


Figure 7.1: Scatter plots of the strategic importance scores for 2000 (black solid dots) and 1991 (grey dots). Again, dots towards the center of the coordinate system are contested units. The plot shows that the number of contested units in 2000 has decreased significantly compared to 1991.

For a test of H1, I compare the overall distribution of ethnic contestation scores before the war to the one after the war. According to H1, post-war ethnic contestation should be significantly lower than before the war. Figure 7.2 shows the empirical distributions of the contestation scores before (solid line) and after the war (dashed line). For the sake of clarity, the distributions are smoothed by a Gaussian kernel density estimator. The figure shows that ethnic contestation in the system decreases during conflict, since the distribution of pre-war scores (solid line) includes higher values than the post-war distribution (dashed line). The result of the visual inspection is confirmed by a numerical comparison. The pre-war mean of contestation is 0.028, as compared to a post-war mean of 0.018. A paired t-test shows that the post-war scores

are significantly lower than pre-war (t-value 7.88, df 108, p-value 0.00).

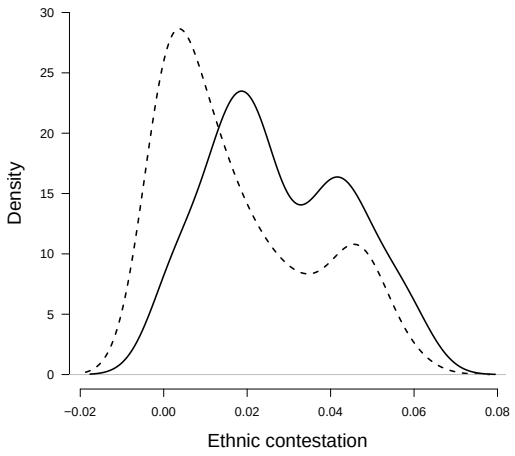


Figure 7.2: Kernel density estimates for the pre-war (solid line) and post-war (dashed line) distribution of ethnic contestation in Bosnia. Post-war ethnic contestation scores are significantly lower.

7.2.2 Local Violence and Decreasing Ethnic Contestation

In the previous paragraph, I showed that during the course of the war, the overall ethnic contestation in the system decreased significantly. How does this change at the systemic level relate to the local occurrence of violence? In what follows, I conduct an analysis of the relationship between one- and two-sided violence and the change in the ethnic contestation of a province. More precisely, the dependent variable is the difference between the 1991 and 2000 scores of ethnic contestation, as measured by the indicator introduced above. This variable ranges from -0.02 to 0.042, with a mean of 0.01. Clearly, as shown in the previous paragraph, during conflict there is a general trend towards lower scores of ethnic contestation, which explains the (on average) positive difference between the 1991 and 2000 scores. The main independent variables are the number of events of one-sided and two-sided violence and are included in a square root transformation because of their highly skewed distribution. Two-sided violence is expected to have a negative effect; in other words, a high number of two-sided violence events should be related to a lower decrease in ethnic contestation (H2). According to H3, we expect one-sided violence to have a positive effect.

I start again with a visual inspection of the bivariate relationship between violence and the change in territorial contestation. For a test of H2, I plot the dependent variable – decrease in ethnic contestation – against the occurrence of two-sided violence. The map is shown in Figure

7.3. For H2 to receive empirical support, two-sided conflict events should be primarily located in areas with low decreases in contestation (light colors).

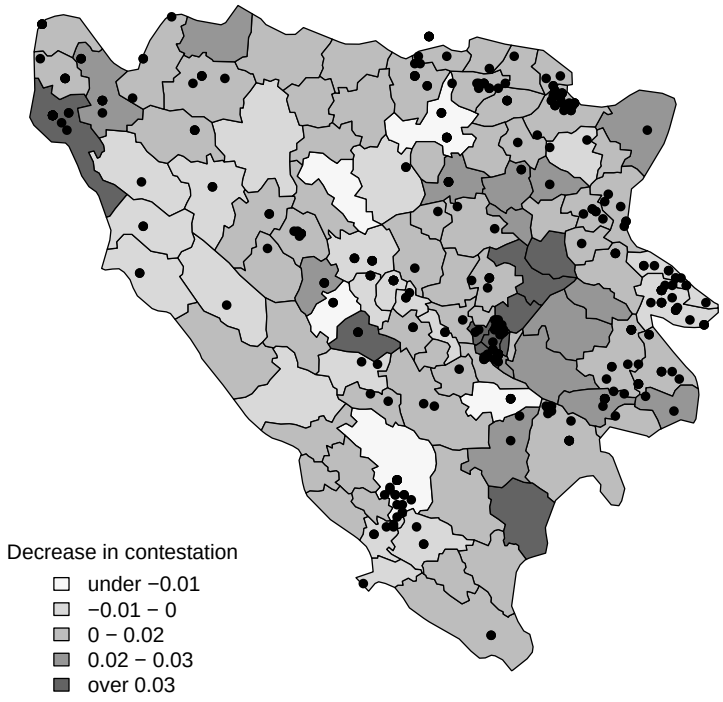


Figure 7.3: Decrease in territorial ethnic contestation between 1991 and 2000. Darker colors correspond to higher decreases in contestation, i.e. greater degrees of strategic unmixing. The dots indicate the location of two-sided conflict events from ACLED.

Figure 7.3 does not reveal a clear pattern. Particularly in the center of the map, there are many areas which seem to be consistent with the proposed relationship. However, we also see a number of counter-examples, such as the Bihac municipality in West Bosnia, or Sarajewo (dense set of points slightly to the east of the map center), which experienced much two-sided violence, but also significant unmixing. In the following, I proceed to a more thorough test using regression analysis. The control variables are largely the same as in the analysis presented in the previous chapter. As shown above, the system moves towards a lower degree of mixing, so more contested units should experience more drastic decreases in ethnic competition. I control for this by including the 1991 level of contestation as an independent variable. Border units should see greater changes, because their strategic position favors their inclusion in the outside group's territory. Lastly, I expect the population of a unit to be be negatively related to the degree of change: In populous units, it should be more difficult to induce changes in the proportion of groups, because many more people need to be "moved" as compared to a unit with a small population. Again, the observations of the dependent variable are likely to

be spatially correlated. Rather than using an OLS model with a spatial lag, I employ a spatial simultaneous autoregressive lag model (Ward and Gleditsch 2008, p. 43) that takes the simultaneity of observations into account.[1] In this model, ρ is the coefficient of the spatially lagged dependent variable. The number of one-sided and two-sided events is highly correlated in the sample (0.73), so I test the two independent variables in separate models. Table 7.1 reports the results with two-sided violence as an independent variable.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	−0.45	1.55	−0.29	0.77
Two-sided violence (sqrt)	−0.17	0.09	−1.80	0.07
Pre-war ethnic contestation	21.21	7.25	2.92	0.00
Border unit	0.00	0.24	0.02	0.98
Population (logged)	0.07	0.15	0.44	0.66
ρ	0.43	0.12	3.76	0.00

Table 7.1: Spatial simultaneous autoregressive lag model. Dependent variable: Decrease in ethnic contestation 1991-2000, multiplied by 100 to ensure better readability (N=109).

In line with H2, two-sided violence leads to a smaller decrease in ethnic contestation in a municipality (Table 7.1): Units with more military confrontations tend to decrease less in their ethnic contestation score. However, the effect is not very strong: Increasing the two-sided event count by one produces a decrease in contestation of .001. This number corresponds to only 3% of the empirical range of the contestation variable. Not surprisingly, pre-war ethnic contestation is strongly and positively associated with the amount of decrease. This is a consequence of the general tendency towards lower ethnic contestation in the system, which causes the largest drops in contestation to occur in the units with high pre-war scores. Border units do not seem to decrease more in contestation, and there is also no discernible effect of the unit population. We see that changes in the ethnic makeup are highly spatially dependent; the coefficient ρ of the spatial lag is positive and strongly significant.

I now repeat this analysis for one-sided violence as the independent variable, starting again with a visual inspection of the data. Figure 7.4 shows a map of the observed decreases in territorial contestation by municipality, similar to Figure 7.3 above. However, the dots in this figure denote reported events of one-sided violence. According to H3, we should see large changes (dark shading) occur at locations with a high number of one-sided events.

Again, it is difficult to discern a clear pattern in Figure 7.4. Many of the municipalities that saw significant changes in territorial contestation (such as the ones in the north or east of

[1] Models estimated using R 2.7.1 and the *spdep* package.

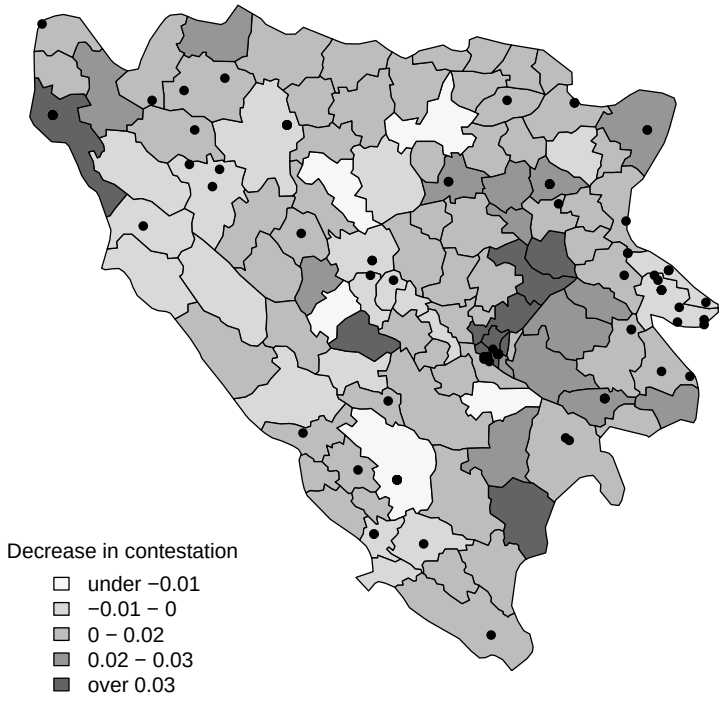


Figure 7.4: Plot of the decrease in ethnic contestation, overlaid with the location of one-sided conflict events from ACLED.

Sarajewo) did not experience one-sided violence at all. However, in some cases the occurrence of one-sided violence did not lead to only marginal contestation decreases, for example in the south of the country. This preliminary analysis does not provide empirical support for H3. Still, it remains to be seen how the relationship is estimated by a regression model. Table 7.2 reports the results for a model with the same specification as above, but with one-sided violence as the independent variable.

The results in Table 7.2 provide no empirical support for my final hypothesis. Contrary to H3, the effect of one-sided violence is negative and not significant. The control variables and the spatial lag retain their signs and significance levels as compared to the previous model. In general, the results suggest that one-sided violence has no direct impact on the triggering of population changes on the ground. Why is this the case? One answer explanation might result from the great differences in the strategies that can be used for ethnic cleansing. As Mann (2005, p. 12) shows, types of cleansing range from active discrimination and cultural repression to violent strategies such as forced replacement, pogroms or genocide. However, in the analysis presented above, we fail to observe these attempts unless they reach a level of violence that classifies them as one-sided violence, and thus qualifies them for inclusion in ACLED. If the

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	−0.20	1.66	−0.12	0.92
One-sided violence (sqrt)	−0.13	0.14	−0.91	0.36
Pre-war ethnic contestation	18.68	7.13	2.62	0.01
Border unit	−0.05	0.24	−0.21	0.83
Population (logged)	0.03	0.16	0.22	0.83
ρ	0.43	0.12	3.71	0.00

Table 7.2: Spatial simultaneous autoregressive lag model. Dependent variable: Decrease in ethnic contestation 1991-2000, multiplied by 100 to ensure better readability (N=109).

majority of population changes in Bosnia was indeed induced by non-violent attempts, it is no surprise that one-sided violence has no effect in the above model. A second alternative explanation is found in the more general pattern of population transfers during the Bosnian war. In my theoretical discussion above, I made the simplifying assumption that people generally tend to stay where they live, unless threatened by violence. In reality, however, this might be different. Burg and Shoup (1999, p. 172) mention completely opposite approaches that leaders pursued during the Bosnia war. Whereas Muslim leaders encouraged their group to stay in their villages in order to maintain the integrity of Bosnia, Serb and Croat leaders tried to convince people to leave their homes and migrate to ethnically homogenous areas. This was done in an attempt to make a future partition of the country more likely. As a result, however, my analysis fails to identity the latter cases, since migration was induced by reasons other than violence.

7.3 Chapter Summary

Conflict changes ethnic settlement patterns, and it does so in a systematic way. The effect of conflict is such that it decreases the level of contestation across all units, corresponding to a pattern of strategic ethnic unmixing during times of war. Whereas frequent confrontations between military forces at a unit seem to be associated with lower degrees of segregation, one-sided violence has not discernible effect on the ethnic map. Rather, the trend towards ethnic homogeneity occurs as a by-product of conflict, but does not seem to be directly induced by the local application of violence. In summary, violence and territory in internal conflict are endogenous: Fuzzy territorial boundaries explain where fighting occurs, and conflict in general leads to an adjustment of these boundaries.

These findings support Kaufmann's (1996) observation that during ethnic conflict, we frequently observe an ethnic unmixing of the population. And still, this should not lead us to accept

the territorial separation of groups as a remedy to conflict (Sambanis 2000; Laitin 2004). This chapter gives at least two additional reasons why this solution might not work. First, territorial separation requires that groups be physically separated. If the resulting configuration should be peaceful, this requires a territorial division accepted by all groups. However, how could such a division be found, and who is to establish it in the first place? The results presented here point to the importance of inter-ethnic boundaries; people are willing to fight for even minimal adjustments. Therefore, an artificial territorial division is unlikely to be agreed on by all groups in a peaceful process. Second, my analysis explains the *location* of violence, given that a civil war has already started. The onset of the civil war itself might be unrelated to the degree of inter-ethnic mixing. As long as we do not know the determinants of civil war onset, it is difficult to estimate the importance of settlement pattern adjustments needed to prevent conflict in the first place. If the underlying conflict issue is not addressed, the drawing of new territorial boundaries will not have the desired effect.

Chapter 8

Conclusion

TERRITORY is important for people – so much that it continues to spur violent conflict around the globe. Building on previous work that establishes the role of group geography in ethnic conflicts, this dissertation makes an attempt to push this research agenda forward by focusing on the mechanisms linking geography and conflict. In other words, the overarching research question of this dissertation is not *whether* group geography influences conflict, but rather *how* it does so.

The purpose of this final chapter is threefold. First, I summarize the key contributions of this dissertation. Referring again to the core questions introduced in Chapter 1, I outline the specific issues and the answers that my analysis gives. Second, I discuss methodological contributions in the study of geography and conflict that might be useful to future research in the field. Third, I discuss policy implications of this work, focusing on geographic risk assessments and territorial solutions to ethnic conflicts.

8.1 Meta-theoretical Contribution

As outlined in Chapter 2, research on civil war in general, but also on group settlement patterns and conflict, suffers from an insufficient focus on the mechanisms behind these phenomena. According to Sambanis (2004*a*), this shortcoming is due to an over-aggregated research design, with a focus on *factors* rather than *actors*. By testing the effect of settlement patterns using coarse indicators, existing studies fail to give a sufficiently precise account of the group processes linking the spatial distribution of groups to conflict involvement. A second category of research – case studies – has produced rich, detailed accounts of these processes, for example regarding questions of participation (Weinstein 2007) or spatial variation in the application of violence (Kalyvas 2006). However, although this literature offers great insights into particular

conflicts, case studies can only offer results of limited generality.

In this dissertation, I attempt to get closer to the group processes during conflict by using a disaggregation approach. Disaggregated studies address the shortcomings of aggregated country- or group-level studies by employing a much more detailed level of analysis, while retaining the explanatory power of large-N comparisons. As elaborated in Chapter 2, disaggregation can be done in a number of ways: conceptually, temporally, and spatially. The analysis presented in the dissertation is an example of the latter type. By using new data and methods in the study of settlement patterns, my work helps to disentangle the group processes during conflict, and thus to fill the gap that has been left by pure correlational studies.

8.2 Theoretical Contributions

This dissertation is built around three core question. I first adopt a focus on a single group and examine how its settlement pattern affects the risk of conflict. The second core question extends this beyond the level of single groups. Here, I narrow the focus down to a detailed geographic level and study how group settlement areas in relation to each other affect the location of violence. Lastly, under the third core question I drop the assumption of territorial configurations as exogenous to conflict, and provide a detailed analysis of how settlement pattern changes relate to violence. In the following paragraphs, I proceed by summarizing the key theoretical contributions of this dissertation for each of the core questions.

8.2.1 Settlement Patterns and the Risk of Conflict

Towards Answering the Motivation vs. Opportunity Question The first core question of this dissertation (presented in Chapters 4 and 5) focuses on how the spatial concentration of a single group affects the risk of being involved in violent conflict with the state. In doing so, the chapter provides a new assessment of the motivation vs.opportunity controversy at the territorial level, which has so far attracted a lot of attention. An example is David Laitin's (2004) article on ethnic unmixing, where he claims that what drives the relationship between group concentration and conflict is the "low-cost communication among the population" of a concentrated group (p. 365). What Laitin advocates here is an opportunity explanation for the effect of group geography on conflict. This perspective, however, stands in contrast to a motivational link from concentration to conflict, which emphasizes the intrinsic value of territory for groups as a determinant of violence (Toft 2003). Until now, no systematic attempt has been made to determine which of these competing explanations bears out empirically. Using novel GIS data on ethnic groups, Chapter 4 presents for the first time a thorough test of both mechanisms. The

results provide evidence in favor of an opportunity-driven influence of group concentration and therefore support Laitin's conjecture. Whereas the relative impact of motivation vs. opportunity determinants of civil war at different levels will most likely continue to be debated, the results presented here at least suggest a clear answer regarding the role of group settlement patterns.

Addressing Group Mobilization in Ethnic Conflicts Based on the results presented in Chapter 4, it becomes clear that a careful study of group mobilization in the context of geography is the way forward. This result calls for even more integration of research on ethnic conflict and collective violence – two strands of literature that have rarely intersected to date (Brubaker 2004, p. 90f). Clearly, the account presented in Chapter 4 is still incomplete in the sense that an opportunity mechanism linking group geography to conflict would necessarily have to specify the group actors involved in conflict mobilization as well as their interactions.

Chapter 5 aims to do this. Using a computational mobilization model run on a real group geography, I simulate how collective mobilization spreads through group populations. As outlined, the literature on group mobilization has repeatedly emphasized the importance of direct social interaction for participation: First, social interaction fosters the socialization of individuals to particular norms and values; second, it facilitates the spread of information; and third, it enables social pressure between group members. Consequently, in my model the influence from one location to the other is proportional to the number of people traveling between these locations, as approximated by Zipf's law (Zipf 1946). The model computes a measure for the simulated pervasiveness of mobilization. More precisely, this indicator captures the proportion of the group that can be mobilized given the parameters of the model, and thus helps to compare groups with respect to their mobilization propensity. Groups with high mobilization values should thus be more likely to be involved in conflict, a result which I confirmed in a regression analysis.

8.2.2 Settlement Patterns and the Location of Violence

Linking the Location of Violence to Settlement Patterns The impact of group settlement patterns on conflict has so far only been examined at the group level. Similar to the results I presented in Chapters 4 and 5, the normal approach has been to study how particular features of group geography affect the group's risk of conflict involvement. The second core question of this dissertation adds a new dimension to the study of group settlement patterns and examines how group geography affects the location of violence. Recent research has made much progress in identifying regions that are particularly prone to conflict, but has so far omitted the spatial ethnic distribution as a key determinant. By employing high-resolution geographic data

on ethnicity and spatially linked reports of violent events in Bosnia, the analysis presented in Chapter 6 sheds new light on the dynamics of civil wars. I show that the application of violence is driven by the territorial incentives of groups to create and consolidate their group territories, i.e. larger areas of ethnic homogeneity. My analysis provides a detailed perspective on different types of ethnic violence by distinguishing between two-sided violence between the groups armed factions and one-sided violence against civilians, both of which are strongly influenced by territorial ambitions.

8.2.3 The Endogeneity of Settlement Patterns and Conflict

Providing a Dynamic Perspective on Group Geography The literature on group settlement patterns and conflict has treated group geography as an exogenous determinant of conflict, which is also the approach in Chapters 4 to 6 of this dissertation. Basically, in these chapters I depart from the assumption that a group's distribution in space is causally prior to the occurrence of conflict. However, the study of conflict at the international level tells us that geography is also the result of armed conflict, or in Tilly's words, "War makes states, and vice versa" (Tilly 1992, p. 67). Chapter 7 applies this thinking to an intrastate setting and analyzes changes in group geography that occur as a result of conflict. Specifically, I pose two questions. First, if violence is driven by competing territorial aspirations of groups, what is the effect of conflict on the level of territorial contestation in general? Second, how do these changes relate to the application of violence on the ground? The answer to the first question comes out clearly in the analysis. The data show that during the conflict in Bosnia, strategic ethnic unmixing occurred. In other words, conflict reduced the territorial contestation that in Chapter 6 was found to drive the application of violence. However, my analysis can only partly explain how local violence relates to these systemic changes. Whereas the occurrence of military confrontations seems to inhibit large changes towards ethnic homogeneity, there is no direct relationship between the targeting of civilians and ethnic unmixing. Instead, population changes seem to be triggered by social and political pressure, which does not manifest itself in violence and thus escapes my analysis.

8.3 Methodological Contributions

Beyond the theoretical contributions outlined in the previous paragraphs, this dissertation makes two major methodological contributions to the study of ethnic conflict, which I will briefly discuss here.

8.3.1 GIS Data on Ethnic Groups

In Political Science, data collected for statistical assessment do not usually measure spatial aspects, and therefore can be adequately represented in a table. Group settlement patterns, however, are inherently spatial characteristics of groups and thus difficult to store in the usual tabular form. As discussed in Chapter 3, numeric indicators of settlement patterns used in previous studies capture the nature of the phenomenon only partially (Toft 2003; Minorities at Risk Project 2005). For this reason, I applied a variety of tools and techniques from Geography to my research question. Most importantly, throughout this dissertation I rely extensively on Geographic Information Systems (GIS), a type of computer software tailored to the collection and processing of spatial data. GIS are only slowly being recognized as a valuable addition to the social science toolkit, but the increased availability of spatial data over the coming years is likely to boost their importance in the field. For this dissertation, GIS-based analysis allows me to distinguish between different aspects of spatial concentration (Chapter 4), but also to prepare data for the computer simulation (Chapter 5) and the analysis of the Bosnia conflict (Chapters 6 and 7). In doing so, this dissertation illustrates how the "spatial turn" in the social sciences can yield precise and testable propositions.

8.3.2 Computer Simulation on Spatial Data

Computational modeling certainly does not belong to the traditional Political Science toolbox, but has gained some momentum over the last decade (Cederman 1997, 2001). However, many agent-based models are used for heuristic purposes and lack any systematic empirical validation. This is certainly one of the reasons why some scholars have remained skeptical about the method in general. In Chapter 5, I addressed this problem by presenting a computational model that is tightly linked to geographic data. Although a few models exist that operate on a real-world geography (see e.g. Dean et al. 2006), my model is unique in that it combines the simulation approach with a large-N statistical test. By using the output of the model as an independent variable in a regression analysis, I show how computational modeling can successfully be applied to a real geography, but at the same time allow for a systematic comparison of the model outcomes to empirical results.

8.4 Policy Implications

What are the practical implications of the research presented in this dissertation? This section discusses two potential policy-related aspects. The first centers around the idea of using settle-

ment patterns as a way to derive risk assessments in a geographic sense, i.e.to identify locations that are at a particularly high risk of being affected by ethnic conflict. Essentially, the gist of this idea lies in its potential to limit the damage during an ethnic conflict, once a war has started. The second application discusses territorial solutions to ethnic wars, as proposed in the literature. Here, the focus is on the termination or prevention of ethnic wars, rather than the precise location of violence. However, as we will see below, my results advise caution with respect to these recommendations.

8.4.1 Geographic Risk Assessments

My analysis on Bosnia presented in Chapter 6 shows that the application of violence follows a territorial pattern, influenced by the groups' strategic objectives. Once a civil war starts, it becomes feasible to predict the spatial progress of conflict. Which municipalities are at a particularly high risk of being affected by violence? Oftentimes, this is the question that third-party military operations face when making decisions about where to intervene (Dorussen 2007). During these operations, protecting the civilian populations might be of particular priority. My results help to single out locations where non-combatants are most likely to be attacked, and can therefore provide guidance as to precisely where precautionary measures would be helpful.

However, for this application one would need a more careful assessment of the predictive performance of the statistical model employed in Chapter 6. Traditional models of civil wars have been found to perform poorly for prediction, despite containing a number of statistically significant variables (Ward and Bakke 2005). Rather than looking at statistical significance as a criterion for the theoretical relevance of independent variables, what is required is an evaluation of the variables and the overall model with respect to predictive accuracy, in other words, their practical relevance. Also, in order to better understand the diffusion processes of violence in ethnic conflicts, it is beneficial to move from a cross-sectional design to a more detailed time-series cross-sectional setup, as demonstrated in Weidmann and Ward (2008).

8.4.2 Territorial Solutions to Civil Wars?

In order for scientific results to resonate within the community of policy-makers, they must encompass factors that can actively be changed from the outside. For example, research on the success of UN peacekeeping operations (Doyle and Sambanis 2000) is highly relevant to policy-makers since it addresses an independent variable (interventions) that can be directly manipulated in order to maximize the desired outcome (in this case, prevent the recurrence of conflict). However, does this apply also to group geography as a conflict determinant? As I

have shown in this dissertation, settlement patterns strongly influence the conflict involvement of particular groups, but also the location where violence is likely to occur. Consequently, is it not advisable to alter the settlement pattern so as to reduce the risk of conflict?

Previous work on the solution of ethnic conflict often treats group geography as a variable that can be tweaked as to make conflict less likely. Most prominently, Kaufmann (1996) advocates the territorial separation of groups as a *geographic* solution to conflict. In particular, he suggests that in order to terminate violent ethnic conflict such as the one in former Yugoslavia, it is necessary to separate group populations into clearly demarcated territories. This suggestion is problematic for many reasons. First and foremost, one might wonder how such a territorial partition could be imposed in the first place. Clearly, this form of "legitimate" ethnic cleansing would entail dramatic costs of human suffering. Furthermore, Sambanis (2000) shows that there is no empirical evidence that supports the proposed relationship between territorial partition and civil war, namely that partition prevents civil wars from recurring. Laitin (2004), on the other hand, points to the security problems caused by the creation of concentrated populations. My results provide strong support for this latter assertion. I show that a high population concentration alleviates collective action problems, thus facilitating group mobilization for conflict. An artificial separation of groups creates a settlement pattern that has precisely that feature and might therefore make things worse in the end. Rather than offer a new solution to the problem of ethnic violence, my results strongly advise against artificial changes of settlement patterns. Thus, other solutions are more likely to succeed in mitigating ethnic violence. Rather than addressing the problem by removing opportunities for conflict, more effective solutions should focus on the motivations groups have for engaging in violence. For example, this motivation can be contained by increasing the autonomy of minorities at the political or cultural level (Lijphart 1977; Hechter 2000*b*). In reducing the fear of majority suppression, this measure might alleviate the motivation for a group to take arms (Hartzell, Hoddie and Rothchild 2001) and is thus better suited to tackle the root cause of violence.

Bibliography

Alesina, Alberto, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat and Romain Wacziarg. 2003. "Fractionalization." *Journal of Economic Growth* 8:155–194.

Anderson, Benedict. 1991. *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. London: Verso.

Andrews, Kenneth T. 2006. "The Dynamics of Protest Diffusion: Movement Organizations, Social Networks, and News Media in the 1960s Sit-Ins." *American Sociological Review* 71:752–777.

Barth, Fredrik. 1969. Introduction. In *Ethnic Groups and Boundaries: The Social Organization of Culture Difference*, ed. Fredrik Barth. Oslo: Universitetsforlaget.

Bates, Robert. 1974. "Ethnic Competition and Modernization in Contemporary Africa." *Comparative Political Studies* 6:457–83.

Beck, Nathaniel, Johnathan N. Katz and Richard Tucker. 1998. "Taking Time Seriously: Time-Series-Cross-Section Analysis with a Binary Dependent Variable." *American Journal of Political Science* 42(4):1260–1988.

Bell-Fialkoff, Andrew. 1993. "A Brief History of Ethnic Cleansing." *Foreign Affairs* 72(3):110–121.

Bhavnani, Ravi. 2006. "Ethnic Norms and Interethnic Violence: Accounting for Mass Participation in the Rwandan Genocide." *Journal of Peace Research* 43(6):651–669.

Biggs, Michael. 2005. "Strikes as Forest Fires: Chicago and Paris in the Late Nineteenth Century." *American Journal of Sociology* 110(6):1684–1714.

Boulding, Kenneth E. 1962. *Conflict and Defense. A General Theory*. New York: Harper and Row.

Bourdieu, Pierre. 1984. "Espace social et genèse des classes." *Actes de la recherche en sciences sociales* 52-53.

Bousquet, F., O. Barreteau, Le Page C., Mullon C. and Weber J. 1999. An Environmental Modelling Approach: The Use of Multi-agent Simulations. In *Advances in Environmental and Ecological Modelling*, ed. F. Blasco. Paris: Elsevier pp. 113–122.

Brass, Paul R. 2003. *The Production of Hindu-Muslim Violence in Contemporary India*. Seattle: University of Washington Press.

Bremer, Stuart A. and Michael Mihalka. 1977. Machiavelli in Machina: Or Politics among Hexagons. In *Problems of World Modeling*, ed. Karl W. Deutsch. Boston: Ballinger pp. 303–337.

Brubaker, Rogers. 2004. *Ethnicity without Groups*. Cambridge: Harvard University Press.

Brubaker, Rogers and Frederick Cooper. 2000. "Beyond "identity"." *Theory and Society* 29(1):1–47.

Bruk, S. I. and V. S. Apenchenko, eds. 1964. *Atlas narodov mira (Atlas of the Peoples of the World)*. Moscow: Glavnoe Upravlenie Geodezii I Kartografi.

Buhaug, Halvard. 2006. "Relative Capability and Rebel Objective in Civil War." *Journal of Peace Research* 43(6).

Buhaug, Halvard and Jan Ketil Rød. 2006. "Local Determinants of African Civil Wars, 1970-2001." *Political Geography* 25(3):315–335.

Buhaug, Halvard, Lars-Erik Cederman and Jan Ketil Rød. 2008. "Disaggregating Ethnic Conflict: A Dyadic Model of Exclusion Theory." *International Organization* 62(3):531–551.

Buhaug, Halvard and Nils Petter Gleditsch. 2006. The Death of Distance? The Globalization of Armed Conflict. In *Territoriality and Conflict in an Era of Globalization*, ed. Miles Kahler and Barbara Walter. Cambridge: Cambridge University Press pp. 187–216.

Buhaug, Halvard and Scott Gates. 2002. "The Geography of Civil War." *Journal of Peace Research* 39(4):417–433.

Burg, Steven L. and Paul S. Shoup. 1999. *The War in Bosnia-Herzegovina: Ethnic Conflict and International Intervention*. New York: M. E. Sharpe, Inc.

Caselli, Francesco and Wilbur John Coleman. 2006. On the Theory of Ethnic Conflict. In *CEP Discussion Papers*. Vol. 732. Available at SSRN: http://ssrn.com/abstract=914166.

Caspersen, Nina. 2004. "Good Fences Make Good Neighbors? A Comparison of Conflict-Regulation Strategies in Postwar Bosnia." *Journal of Peace Research* 41(5):569–588.

Cederman, Lars-Erik. 1997. *Emergent Actors in World Politics: How States and Nations Develop and Dissolve*. Princeton, NJ: Princeton University Press.

Cederman, Lars-Erik. 2001. "Agent-Based Modeling in Political Science." *The Political Methodologist* 10:16–22.

Cederman, Lars-Erik. 2002*a*. "Endogenizing Geopolitical Boundaries with Agent-based Modeling." *Proceedings of the National Academy of Sciences USA* 99(3):7296–7303.

Cederman, Lars-Erik. 2002*b*. Nationalism and Ethnicity. In *Handbook of International Relations*, ed. Walter Carlsnaes, Thomas Risse and Beth Simmons. London: Sage.

Cederman, Lars-Erik. 2003. "Modeling the Size of Wars: From Billiard Balls to Sandpiles." *American Political Science Review* 97(1):135–150.

Cederman, Lars-Erik and Christopher Daase. 2003. "Endogenizing Corporate Identities: The Next Step in Constructivist IR Theory." *European Journal of International Relations* 9(1):5–35.

Cederman, Lars-Erik and Luc Girardin. 2007. "Beyond Fractionalization: Mapping Ethnicity onto Nationalist Insurgencies." *American Political Science Review* 101(01):173–185.

Center for International Earth Science Information Network CIESIN, Columbia University; and Centro Internacional de Agricultura Tropical CIAT. 2005. "Gridded Population of the World v3 (GPWv3).". Available at http://sedac.ciesin.columbia.edu/gpw.

Chandra, Kanchan. 2001. "Introduction: Constructivist Findings and Their Non-Incorporation." *APSA-CP: Newsletter of the Organized Section in Comparative Politics of the American Political Science Association* 12(1):7–11.

Chandra, Kanchan. 2006. "What is Ethnic Identity and Does It Matter?" *Annu. Rev. Polit. Sci.* 9:397–424.

Checkel, Jeffrey T. 2008. "Causal Mechanisms and Civil War." Paper prepared for the Annual Meeting of the American Political Science Association, Boston, MA.

Chwe, Michael Suk-Young. 1999. "Structure and Strategy in Collective Action." *American Journal of Sociology* 105(1):128–156.

Coleman, James S. 1990. *Foundations of Social Theory*. Cambridge, MA: The Belknap Press of Harvard University Press.

Collier, Paul. 2003. *Breaking the Conflict Trap: Civil War and Development Policy*. Oxford University Press.

Collier, Paul and Anke Hoeffler. 2004. "Greed and Grievance in Civil War." *Oxford Economic Papers* 56(4):563–595.

Cornell, Svante E. 2002. "Autonomy as a Source of Conflict: Caucasian Conflicts in Theoretical Perspective." *World Politics* 54:245–76.

de Figueiredo, Rui and Barry R. Weingast. 1999. The Rationality of Fear: Political Opportunism and Ethnic Conflict. In *Civil Wars, Insecurity, and Intervention*, ed. Barbara F Walter and Jack Snyder. New York: Columbia University Press.

Dean, Jeffrey S., George J. Gumerman, Joshua M. Epstein, Robert Axtell, Alan C. Swedlund, Miles T. Parker and Stephen McCarroll. 2006. Understanding Anasazi Culture Change through Agent-based Modeling. In *Generative Social Science: Studies in Agent-based Computational Modeling*, ed. Joshua M. Epstein. Princeton: Princeton University Press.

Diani, Mario. 2003. Introduction. In *Social Movements and Networks: Relational Approaches to Collective Action*, ed. Mario Diani and Doug McAdam. Oxford: Oxford University Press.

Diehl, Paul F. 1991. "Geography and War: A Review and Assessment of the Empirical Literature." *International Interactions* 17(1):11–27.

Dollard, J., L. W. Doob, N. E. MIller, O. H. Mowrer and R. R. Sears. 1939. *Frustration and Aggression*. New Haven: Yale University Freer.

Dorussen, Han. 2007. "Introducing PKOLED: A Peacekeeping Operations Location and Event Dataset." Paper prepared for the Conference on Disaggregating the Study of Civil War and Transnational Violence, University of Essex, UK.

Doyle, Michael W. and Nicholas Sambanis. 2000. "International Peacebuilding: A Theoretical and Quantitative Analysis." *American Political Science Review* 94(4).

Epstein, Joshua M. 2006. *Generative Social Science: Studies in Agent-based Computational Modeling*. Princeton: Princeton University Press.

ESRI. 1998. "ESRI Shapefile Technical Description.". Available at http://esri.com/library/whitepapers/pdfs/shapefile.pdf.

Fearon, James D. 1995. "Rationalist Explanations for War." *International Organization* 49(3):379–414.

Fearon, James D. 1998. Commitment Problems and the Spread of Ethnic Conflict. In *The International Spread of Ethnic Conflict: Fear, Diffusion and Escalation*, ed. David A. Lake and Donald Rothchild. Princeton: Princeton University Press.

Fearon, James D. 2003. "Ethnic and Cultural Diversity by Country." *Journal of Economic Growth* 8:195–222.

Fearon, James D. 2004. "Why Do Some Civil Wars Last So Much Longer Than Others?" *Journal of Peace Research* 41(3):275–301.

Fearon, James D. 2006. Ethnic Mobilization and Ethnic Violence. In *The Oxford Handbook of Political Economy*, ed. Barry R. Weingast and Donald A. Wittman. New York: Oxford University Press.

Fearon, James D. and David D. Laitin. 1996. "Explaining Interethnic Cooperation." *American Political Science Review* 90(4).

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

Fearon, James D. and David D. Laitin. 2006. "Lithuania.". Typescript, Stanford University. Available at http://www.stanford.edu/group/ethnic/Random Narratives/LithuaniaRN1.3.pdf.

Fox, Jonathan. 2002. "Ethnic Minorities and the Clash of Civilizations: A Quantitative Analysis of Huntington's Thesis." *British Journal of Political Science* 32(3):415–434.

Frigg, Roman and Stephan Hartmann. 2006. Models in Science. In *Stanford Encyclopedia of Philosophy*. Available at http://plato.stanford.edu/entries/models-science/.

Gagnon, V. P. 1995. "Ethnic Nationalism and International Conflict: The Case of Serbia." *International Security* 19(3).

Gagnon, V. P. 2004. *The Myth of Ethnic War: Serbia and Croatia in the 1990s*. Ithaca: Cornell University Press.

Gartzke, Erik and Kristian Skrede Gleditsch. 2006. "Identity and Conflict: Ties that Bind and Differences that Divide." *European Journal of International Relations* 12(1):53–87.

Gates, Scott. 2002. "Recruitment and Allegiance: The Microfoundations of Rebellion." *Journal of Conflict Resolution* 46(1):111–130.

Geertz, Clifford. 1963. The Integrative Revolution: Primordial Sentiments and Civil Politics in the New States. In *Old Societies and New States*, ed. Clifford Geertz. New York: Free Press.

Gellner, Ernest. 1983. *Nations and Nationalism*. Oxford: Blackwell.

George, Alexander L. and Andrew Bennett. 2004. *Case Studies and Theory Development in the Social Sciences*. Cambridge: MIT Press.

Giddens, Anthony. 1984. *The Constitution of Society. Outline of the Theory of Structuration*. Cambridge: Polity.

Gilbert, Nigel. 2007. *Agent-based Models*. London: Sage.

Gilbert, Nigel and Klaus G. Troitzsch. 2005. *Simulation for the Social Scientist*. 2nd ed. Open University Press.

Gilpin, Robert. 1981. *War and Change in World Politics*. Cambridge: Cambridge University Press.

Gimblett, H. Randy, ed. 2002. *Integrating Geographic Information Systems and Agent-Based Modeling for Simulating Social and Ecological Processes*. Oxford: Oxford University Press.

Gleditsch, Kristian Skrede. 2002. *All International Politics Is Local: The Diffusion of Conflict, Integration and Democratization*. University of Michigan Press.

Gleditsch, Nils Petter, Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg and Håvard Strand. 2002. "Armed Conflict 1946-2001: A New Dataset." *Journal of Peace Research* 39(5):615–637.

Goddard, Stacie E. 2006. "Uncommon Ground: Indivisible Territory and the Politics of Legitimacy." *International Organization* 60:35–68.

Goemans, Hein. 2006. Bounded communities: territoriality, territorial attachment, and conflict. In *Territoriality and Conflict in an Era of Globalization*, ed. Miles Kahler and Barbara Walter. Cambridge: Cambridge University Press.

Goemans, Hein E., Kristian Skrede Gleditsch and Giacomo Chiozza. 2009. "Introducing Archigos: A Data Set of Political Leaders." *Journal of Peace Research* 46(2).

Gordon, Raymond G. Jr, ed. 2005. *Ethnologue: Languages of the World, Fifteenth edition*. Dallas: SIL International. Online version: http://www.ethnologue.com/.

Gourevitch, Philip. 1999. *We Wish to Inform You That Tomorrow We Will be Killed With Our Families: Stories from Rwanda*. Picador.

Granovetter, Mark. 1978. "Threshold Models of Collective Behavior." *American Journal of Sociology* 83(6):1420–1443.

Gurr, Ted. 1968. "A Causal Model of Civil Strife: A Comparative Analysis Using New Indices." *American Political Science Review* 62(4):1104–1124.

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

Gurr, Ted Robert. 1993. *Minorities at Risk: A Global View of Ethnopolitical Conflicts*. Washington, DC: United States Institute of Peace Press.

Hägerstrand, Thorsten. 1967. *Innovation diffusion as a Spatial Process*. Chicago: University of Chicago Press.

Harbom, Lotta, Erik Melander and Peter Wallensteen. 2008. "Dyadic Dimensions of Armed Conflict, 1946-2007." *Journal of Peace Research* 45(5):697–710.

Harbom, Lotta, Håvard Strand and Håvard M. Nygard. 2008. "UCDP/PRIO Armed Conflict Dataset Codebook.". Available at http://www.prio.no/sptrans/1664678440/Codebook.pdf.

Hardin, Russell. 1995. *One for All: The Logic of Group Conflict*. Princeton: Princeton University Press.

Harris, Chauncy D. 1965. "Review of the Atlas Narodov Mira." *Geographical Review* 55(4):608–610.

Hartzell, Caroline, Matthew Hoddie and Donald Rothchild. 2001. "Stabilizing the Peace after Civil War: An Investigation of Some Key Variables." *International Organization* 55(1):183–208.

Hechter, Michael. 2000*a*. *Containing Nationalism*. New York: Oxford University Press.

Hechter, Michael. 2000*b*. "Nationalism and Rationalism." *Journal of World Systems Research* 6(2):308–329.

Hechter, Michael and Dina Okamoto. 2001. "Political Consequences of Minority Group Formation." *Annu. Rev. Polit. Sci.* 4:189–215.

Hedström, Peter and Richard Swedberg. 1998. Social mechanisms: An introductory essay. In *Social Mechanisms: An Analytical Approach to Social Theory*, ed. Peter Hedström and Richard Swedberg. Cambridge: Cambridge University Press.

Hegre, Håvard. 2008. "Gravitating toward War: Preponderance May Pacify, but Power Kills." *Journal of Conflict Resolution* 52(4):566–589.

Hegre, Håvard and Clionadh Raleigh. 2006. "Population Size, Concentration and Civil War: A Geographically Disaggregated Analysis." Paper prepared for the Annual Meeting of the International Studies Association, San Diego, CA.

Hempel, Carl G. 1965. *Aspects of Scientific Explanation*. New York: Free Press.

Hensel, Paul R. 2000. Territory: Theory and Evidence on Geography and Conflict. In *What Do We Know About War?*, ed. John A. Vasquez. Lanham: Rowman and Littlefield.

Herbst, Jeffrey. 2000. *States and Power in Africa: Comparative Lessons in Authority and Control*. Princeton, NJ: Princeton University Press.

Hewes, Gordon W. 1966. "Review of the Atlas Narodov Mira." *American Anthropologist* 68:532–534.

Hoffmann, Matthew J. 2008. Agent-based Modeling. In *Qualitative Methods in International Relations: A Pluralist Guide*, ed. A. Klotz and D. Prakash. Palgrave Macmillan pp. 187–208.

Holsti, Kalevi Jacque. 1991. *Peace and War: Armed Conflicts and International Order, 1648-1989*. Cambridge: Cambridge University Press.

Horowitz, Donald. 1985. *Ethnic Groups in Conflict*. Berkeley: University of California Press.

Horowitz, Donald. 2002. *The Deadly Ethnic Riot*. University of California Press.

Hughes, James and Gwendolyn Sasse. 2002. *Ethnicity and Territory in the Former Soviet Union: Regions in Conflict*. Routledge.

Huntington, Samuel P. 1993. "The Clash of Civilizations?" *Foreign Affairs* 72(3):22–49.

Huntington, Samuel P. 1996. *The Clash of Civilizations and the Remaking of World Order*. New York: Simon and Schuster.

Huth, Paul K. 2000. Territory: Why are Territorial Disputes between States a Central Cause of International Conflict? In *What Do We Know About War?*, ed. John A. Vasquez. Lanham: Rowman and Littlefield.

Imai, Kosuke, Gary King and Olivia Lau. 2006. "Zelig: Everyone's Statistical Software.". http://gking.harvard.edu/zelig/.

International Crisis Group. 2006. "Abkhazia Today." *Europe Report* (176).

Isard, Walter. 1954. "Location Theory and Trade Theory: Short-Run Analysis." *Quarterly Journal of Economics* 68:305–322.

Kalyvas, Stathis. 2006. *The Logic of Violence in Civil War*. New York: Cambridge University Press.

Kaplan, R. D. 1993. *Balkan Ghosts: A Journey through History*. New York: St. Martin's Press.

Kaufmann, Chaim. 1996. "Possible and Impossible Solutions to Ethnic Civil Wars." *International Security* 20(4):136–175.

King, Gary, Robert O Keohane and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.

Kocher, Matthew and Stathis Kalyvas. 2003. "An Analysis of the Hamlet Evaluation Survey (HES)." Paper prepared for the Annual Meeting of the American Political Science Association, Philadelphia, PA.

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

Kuran, Timur. 1991. "Now out of Never: The Element of Surprise in the East European Revolution of 1989." *World Politics* 44:7–48.

Lacina, Bethany and Nils Petter Gleditsch. 2005. "Monitoring Trends in Global Combat: A New Dataset of Battle Deaths." *European Journal of Population* 21(2-3).

Laitin, David D. 2004. "Ethnic Unmixing and Civil War." *Security Studies* 13(4):350–365.

Larsen, Jonas, John Urry and Kay Axhausen. 2006. *Mobilities, Networks, Geographies*. Ashgate Publishing, Ltd.

Latané, Bibb. 1981. "The Psychology of Social Impact." *American Psychologist* 36:343–365.

Lazarsfeld, Paul. 1968. *People's Choice: How the Voter Makes Up His Mind in a Presidential Campaign*. 3rd ed. Columbia University Press.

Lemke, Douglas. 1995. "The Tyranny of Distance: Redefining Relevant Dyads." *International Interactions* 21:23–38.

Levinson, David. 1998. *Ethnic Groups Worldwide: A Ready Reference Handbook*. Phoenix: The Oryx Press.

Lichbach, Mark Irving. 1995. *The Rebel's Dilemma*. University of Michigan Press.

Licklider, Roy. 1995. "The Consequences of Negotiated Settlements in Civil Wars, 1945-1993." *American Political Science Review* 89(3):681–690.

Lijphart, Arend. 1977. *Democracy in Plural Societies*. New Haven: Yale University Press.

Lim, May, Richard Metzler and Yaneer Bar-Yam. 2007. "Global Pattern Formation and Ethnic/Cultural Violence." *Science* 317:1540–1544.

Little, Daniel. 1991. *Varieties Of Social Explanation: An Introduction To The Philosophy Of Social Science*. Westview Press.

Lohmann, Susanne. 1993. "A Signaling Model of Informative and Manipulative Political Action." *American Political Science Review* 87(2):319–333.

Lujala, Päivi, Jan Ketil Rød and Nadja Thieme. 2007. "Fighting over Oil: Introducing a New Dataset." *Conflict Management and Peace Science* 24(3):239–256.

Lujala, Päivi, Nils Petter Gleditsch and Elisabeth Gilmore. 2005. "A Diamond Curse? Civil War and a Lootable Resource." *Journal of Conflict Resolution* 49:538–562.

Mann, Michael. 2005. *The Dark Side of Democracy: Explaining Ethnic Cleansing*. Cambridge University Press.

Marshall, Monty G. and Keith Jaggers. 2000. "Polity IV Dataset and Users' Manual: Political Regime Characteristics and Transitions, 1800-1999.". Available at http://www.cidcm.umd.edu/polity/.

Marshall, Monty G., Ted Gurr and Barbara Harff. 2008. "Internal Wars and Failures of Governance, 1955-2006: Political Instability Task Force Problem Set Codebook.". Available at http://globalpolicy.gmu.edu/pitf/pitfcode.htm.

Mayntz, Renate. 2003. "Mechanisms in the Analysis of Social Macro-Phenomena." *Philosophy in the Social Sciences* 34(2):237–259.

McColl, Robert. 1969. "The Insurgent State: Territorial Bases of Revolution." *Annals of the Association of American Geographers* 59(4):613–631.

Meadows, Donella H., Dennis L. Meadows, Jørgen Randers and William W. Behrens. 1972. *Limits to Growth*. Universe Books.

Melander, Erik. 2001. "The Nagorno-Karabakh Conflict Revisited: Was the War Inevitable?" *Jouirnal of Cold War Studies* 3(2):48–75.

Melander, Erik. 2007. "Ethnic Cleansing in Bosnia-Herzegovina, 1992-1995." Paper prepared for the Conference on Disaggregating the Study of Civil War and Transnational Violence, University of Essex.

Miller, Harvey J. 2004. "Tobler's First Law and Spatial Analysis." *Annals of the Association of American Geographers* 94(2):284–289.

Minorities at Risk Project. 2005. "Minorities at Risk Dataset.". Available at http://www.cidcm.umd.edu/mar/.

Minorities at Risk Project. 2006. "Minorities At Risk Organizational Behavior (MAROB) Provisional Code Book v4.6.". Center for International Development and Conflict Management (CIDCM) and Center for the Study of Terrorism and Responses to Terrorism (START).

Moore, Will. 1995. "Rational Rebels: Overcoming the Free-Rider Problem." *Political Research Quarterly* 48:417–454.

Moss, Scott. 2008. "Alternative Approaches to the Empirical Validation of Agent-Based Models." *Journal of Artificial Societies and Social Simulation* 11(1):5. 1460-7425.

Mueller, John. 2000. "The Banality of "Ethnic War"." *International Security* 25(1):42–70.

Newman, Saul. 1991. "Does Modernization Breed Ethnic Political Conflict?" *World Politics* 43(3):451–478.

Nowak, Andrzej, Jacek Szamrej and Bibb Latané. 1990. "From Private Attitude to Public Opinion: A Dynamic Theory of Social Impact." *Psychological Review* 97(3):362–376.

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

Osborne, Martin J. and Ariel Rubinstein. 1994. *A Course in Game Theory*. Cambridge: The MIT Press.

Petersen, Roger D. 2002. *Understanding Ethnic Violence: Fear, Hatred and Resentment in Twentieth Century Eastern Europe*. New York: Cambridge University Press.

Petrovic, Ruza. 1992. "The National Composition of Yugoslavia's Population, 1991." *Yugoslav Survey* 33(1):3–24.

Pinard, Maurice. 1968. "Mass Society and Political Movements: A New Formulation." *American Journal of Sociology* 73(6):682–690.

Posen, Barry R. 1993*a*. The Security Dilemma and Ethnic Conflict. In *Ethnic Conflict and International Security*, ed. Michael E. Brown. Princeton: Princeton University Press.

Posen, Barry R. 1993*b*. "The Security Dilemma and Ethnic Conflict." *Survival* 35(1):27–47.

Posner, Daniel N. 2004. "Measuring Ethnic Fractionalization in Africa." *American Journal of Political Science* 48(4):849–863.

Raleigh, Clionadh and Håvard Hegre. 2005. "Introducing ACLED: An Armed Conflict Location and Event Dataset." Paper prepared for the Conference on Disaggregating the Study of Civil War and Transnational Violence, San Diego, CA.

Richardson, Lewis Fry. 1961. "The Problem of Contiguity: An Appendix to Statistics of Deadly Quarrels." *General Systems Yearbook* 6:140–187.

Ross, Russell R. 1987. *Cambodia: A Country Study*. Washington: GPO for the Library of Congress.

Salehyan, Idean. 2007. "Refugees and the Study of Civil War." *Civil Wars* 9(2):127–141.

Sambanis, Nicholas. 2000. "Partition as a Solution to Ethnic War." *World Politics* 52:437–483.

Sambanis, Nicholas. 2004*a*. "Using Case Studies to Expand Economic Models of Civil War." *Perspectives on Politics* 2(2):259–279.

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

Sawyer, Keith. 2004. "Social Explanation and Computational Simulation." *Philosophical Explorations* 7(3):219–231.

Schelling, Thomas C. 1971. "Dynamic Models of Segregation." *Journal of Mathematical Sociology* 1:143–186.

Schelling, Thomas C. 1978. *Micromotives and Macrobehavior*. New York: Norton.

Siegfried, André. 1975. *Tableau politique de la France de l'ouest sous la Troisième République: The Political Map of Western France Under the Third Republic*. Ayer Publishing.

Singer, J. David and Melvin Small. 1972. *The Wages of War, 1816-1965: A Statistical Handbook*. New York: Wiley.

Small, Melvin and J. David Singer. 1982. *Resort to Arms: International and Civil Wars*. Beverly Hills, CA: Sage.

Smith, Anthony. 1986. *The Ethnic Origins of Nations*. Oxford: Blackwell Publishers.

Srbljinovic, Armano, Drazen Penzar, Petra Rodik and Kruno Kardov. 2003. "An Agent-based Model of Ethnic Mobilization." *Journal of Artificial Societies and Social Simulation* 6(1).

Starr, H. and G. D. Thomas. 2002. "The 'Nature' of Contiguous Borders: Ease of Interaction, Salience, and the Analysis of Crisis." *International Interactions* 28:213–235.

Starr, Harvey and Benjamin Most. 1978. "A Return Journey: Richardson, Frontiers, and War in the 1945-1965 Era." *Journal of Conflict Resolution* 22:441–462.

Straus, Scott. 2004. "How many perpetrators were there in the Rwandan genocide? An estimate." *Journal of Genocide Research* 6(1):85–98.

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

Taylor, Charles and Michael C. Hudson. 1972. *World Handbook of Political and Social Indicators*. New Haven: Yale University Press.

Tilly, Charles. 1978. *From Mobilization to Revolution*. Addison-Wesley.

Tilly, Charles. 1992. *Coercion, Capital, and European States*. Studies in Social Discontinuity Blackwell Publishing.

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

Tobler, Waldo R. 1970. "A computer movie simulating urban growth in the Detroit region." *Economic Geography* 46:234–240.

Toft, Monica Duffy. 2002. "Indivisible Territory, Geographic Concentration, and Ethnic War." *Security Studies* 12(2):82–119.

Toft, Monica Duffy. 2003. *The Geography of Ethnic Violence: Identity, Interests, and the Indivisibility of Territory*. Princeton, NJ: Princeton University Press.

United Nations High Commissioner for Refugees. 2006. *The State of the World's Refugees*. New York: Oxford University Press.

U.S. Department of State. 1999. "Ethnic Cleansing in Kosovo: An Accounting.". Available at http://www.state.gov/www/global/human_rights/kosovoii/pdf/kosovii.pdf.

Van Evera, Stephen. 2001. "Primordialism Lives!" *APSA-CP: Newsletter of the Organized Section in Comparative Politics of the American Political Science Association* 12(1):20–22.

Varshney, Ashutosh. 2002. *Ethnic Conflict and Civic Life*. Yale University Press.

Waltz, Kenneth N. 1979. *Theory of International Politics*. McGraw-Hill.

Ward, Michael D. and Kristian Skrede Gleditsch. 2008. *Spatial Regression Models*. Sage.

Ward, Michael D. and Kristin Bakke. 2005. "Predicting Civil Conflicts: On the Utility of Empirical Research." Paper prepared for the Conference on Disaggregating the Study of Civil War and Transnational Violence, San Diego, CA.

Warren, T. Camber. 2007. "Communicative Structure and the Emergence of Sociopolitical Fragmentation." Paper prepared for the Annual Meeting of the American Political Science Association, Chicago, IL.

Weidmann, Nils B. 2009. "Kausalprozesse auf mehreren Ebenen: Ein Plädoyer für vertikale Integration in der Bürgerkriegsforschung." *Politische Vierteljahresschrift* . Forthcoming.

Weidmann, Nils B. and Michael D. Ward. 2008. "Spatial-temporal Modeling of Civil War: The Example of Bosnia." Paper prepared for the GROW-Net Conference, Zurich.

Weiner, Myron. 1971. "The Macedonian Syndrome: A Historical Model of International Relations and Political Development." *World Politics* 23(4):665–683.

Weinstein, Jeremy M. 2007. *Inside Rebellion*. New York: Cambridge University Press.

Wendt, Alexander. 1994. "Collective Identity Formation and the International State." *American Political Science Review* 88:384–396.

Wesley, James Paul. 1962. "Frequency of wars and geographical opportunity." *The Journal of Conflict Resolution* 6(4):387–389.

Wilkinson, Steven. 2004. *Votes and Violence*. Cambridge University Press.

Windrum, Paul, Giorgio Fagiolo and Alessio Moneta. 2007. "Empirical Validation of Agent-Based Models: Alternatives and Prospects." *Journal of Artificial Societies and Social Simulation* 10(2):8. 1460-7425.

Witten, Ian and Eibe Frank. 2005. *Data Mining: Practical Machine Learning Tools and Techniques*. 2nd ed. Morgan Kaufmann.

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

Zeilis, Achim, Christian Kleiber and Simon Jackman. 2007. "Regression Models for Count Data in R." *WU Wien Research Report Series* 53.

Zipf, George Kingsley. 1946. "The P1P2/D Hypothesis: On the Intercity Movement of Persons." *American Sociological Review* 11(6):677–686.

Zuckerman, Alan S. 2005. Returning to the Social Logic of Politics. In *The Social Logic of Politics: Personal Networks as Contexts for Political Behavior*, ed. Alan S. Zuckerman. Philadelphia: Temple University Press.