“All Conflict is Local”:

an Empirical Analysis of Local Factors in Violent Civil Conflict

DPhil Thesis in Politics

Whitney Haring-Smith
St. John’s College
Oxford University

"ALL CONFLICT IS LOCAL":

AN EMPIRICAL ANALYSIS OF LOCAL FACTORS IN VIOLENT CIVIL CONFLICT

Abstract

Previous civil war analyses have approached conflict as a single category with limited exceptions, and this thesis project assesses whether differentiating conflicts by their type and intensity using a local-level geo-referenced analytical approach produces differing results for sub-groups of conflicts. The conflicts are divided into 1) governmental hostilities, where the aim of the armed non-state group is to capture the state, and 2) territorial hostilities, where the aim of the armed non-state group is to capture increased autonomy or secession for a territorial claim. The conflicts are also differentiated by intensity into 1) low-intensity conflicts, with fewer than 1000 battle-related deaths per year, and 2) civil wars, with 1000 or more battle-related deaths per year. The results demonstrate that conflicts with differing insurgent goals and intensities of battle are correlated with markedly different factors. There are three factors – local population density, change in local rainfall, and statewide GDP growth – that are significant to both governmental and territorial hostilities but have opposite signs for the two sets. Only one variable – Polity IV scores – showed a consistently significant correlation for governmental and territorial hostilities. There are no factors that are significant to both low-intensity conflict and higher-intensity civil war. These findings suggest that approaching all conflicts as a single class, particularly at the local level, may not reveal significant differences in factors correlated with conflict. Modeling of local conflict will require differentiation of conflicts into salient sub-groups. For policymakers and practitioners, this research suggests that there is not a one-size-fits-all approach for conflict prevention but that strategies need to be targeted to specific types of conflict.

Contact Information
Whitney Haring-Smith
DPhil Candidate in Politics
St. John's College, Oxford
Email: whitney.haring-smith@aya.yale.edu
Email (alternate): whitney.haring-smith@politics.ox.ac.uk
Phone: +1 724 964 4404
Skype: wharings

"ALL CONFLICT IS LOCAL":

AN EMPIRICAL ANALYSIS OF LOCAL FACTORS IN VIOLENT CIVIL CONFLICT

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

Over the past three years, this project has progressed from an idea that political, economic, and geographic data should be combined for studying conflict into a full-fledged dataset with a defined methodological approach for accomplishing the final models. Studying something as raw and passionate as violent conflict through the use of statistics and equations seems, at first glance, to be a futile concept. Over years of study, though, research on civil wars has developed a significant statistical component. Academic work has evolved considerably from theoretical discussions of conflict trends based on rudimentary statistical assessments onward to regressions of decades of country data, and yet further on to disaggregated sub-national assessments. Each of these developments has, in turn, substantially improved our understanding of conflict within countries. This project attempts to refine the existing literature by demonstrating that all conflicts cannot be explained by the same set of analytical tools or factors. Rather, when conflicts are differentiated by their type or intensity, results differ – sometimes in significant and opposing directions – for the role that some factors, such as GDP growth, play in local conflict dynamics. Thus, previous efforts to study conflicts at the state level using one regression to assess all types of civil wars may have missed some important distinctions.

Before engaging the details of the methodological approach, it is important to understand the basis from which existing theory and analytical practice on civil war has been derived. Some of the first theoretical ponderings on the causes of civil war can be traced back to Plato's *Republic* and Aristotle's *Politics* in the Western canon, Confucius and others in the Eastern canon, and likely many more. Mostly philosophical in nature, these works relied on case studies to form the basis of evidence. Over the subsequent centuries, scholars, mostly military, have developed more sophisticated theories for the causes of civil war, but the analysis of most major theorists, such as Machiavelli and Clauswitz, was still driven by case studies and personal experiences. Examinations of civil conflict prior to the middle of the 20th century rarely included any form of quantitative analysis.

In the middle of the 1900s, both military thinkers and academic scholars began to converge on quantitative methods as a way of understanding conflict. While military theorists and practitioners, such as Carl von Clauswitz and T.E. Lawence, used numerous metaphors using the principles of physics to describe armed conflict (Clausewitz 1976; Lawrence 1929), the academic work of Lewis Richardson in the middle of the 1900s appears to be the first clear use of statistics to assess human conflict (Rapoport 1957; Sutherland 1962; Cederman 2003). A physicist by training, it is understandable that Richardson would look towards statistics to develop model of conflict (Richardson 1948, 1960). As Rapoport explains Richardson's thoughts,

> If one grants that physical determinism has limits on the microscopic end of the scale, one has to grant the possibility that the indeterminacy seemingly inherent in situations involving human motivations, volitions, etc., *may* disappear as the events are taken on a larger and larger scale. The trend is away from old mystical, teleological determinism and toward a statistical one" (Rapoport 1957).

While Richardson's work focused mostly on international conflict and other "deadly quarrels," his application of quantitative analysis to conflict laid the groundwork for broader studies in the 1960s and 1970s on both international and civil conflict (Denton

1966; Gurr 1968, 1970; Duvall and Wefling 1973).[1] Singer and Diehl aptly describe the transition to the use of statistical methods, writing,

> Until the "behavioral" movement arrived, research in international politics differed from the study of diplomatic and military history only in that the former was more polysyllabic and conceptually rich and more contemporary in its focus…we pursued knowledge in essentially the same fashion as our colleagues in the history departments. That is, we held and articulated rather vague "theories," rested them on an uncounted number of inarticulate and imprecise assumptions, and let those assumptions shape the stories that we told and the inferences that we drew. It is only in the last thirty years that scholars have begun to conduct systematic, empirical research on international political phenomena. (Singer and Diehl 1990)

As the availability of data and the attention to empirical evidence supporting theories grew, quantitative analysis of conflict became increasingly common in political science and conflict studies.

Even as the early reviews of the progress towards empirical methods began to be written, such as the above quoted review by Singer and Diehl in 1990, empirical methods had little impact on policy formation. Scholars remained focused on issues of how to improve measurement for the data used in empirics and develop a more robust and credible set of findings from statistical methods. Moreover, much of the initial empirical work, particularly from the Correlates of War Project, focused initially on interstate conflict and describing the international system rather than on civil conflicts. The end of the Cold War and the corresponding drop in attention to *inter*state conflict and rise in consideration of

[1] It is interesting to note that some of the earliest cross-national statistical studies of civil conflict, such as Duvall and Wefling's 1973 "Social Mobilization, Political Institutionalization, and Conflict in Black Africa: A Simple Dynamic Model," focused on Africa, as have the early disaggregated and geostatistical studies. This issue will be taken up in greater detail later on, but there has been a small literature on distinctions between Africa and the rest of the world that may make early Africa-only results less representative of conflict trends as a whole (Elbadawi and Sambanis 2000; Collier and Hoeffler 2002).

*intra*state conflict stimulated new research into the causes of civil wars, and heightened policy interest in the results of such research. The World Bank supported the development of empirical studies as a tool to assist the Bank in promoting development in conflict-wracked areas of the world, so World Bank funded the creation and analysis of detailed datasets that tracked variables associated with civil war. Over the 1990s, empirical civil war studies gained credibility and attention, eventually becoming central in academic, policy and public debates. As Fearon noted, Collier and Hoeffler's 2004 study of the relationship between primary commodity exports and civil war garnered broad media coverage and policy traction:

> First posted on the World Bank Web site in early 2000…this finding (Collier and Hoeffler 2004) received widespread press coverage, garnering articles and editorial comment in *The Economist,* the *New York Times*, the *Washington Post*, and the *Financial Times*, among many other newspapers and magazines (Fearon 2005).

The publicity enjoyed statistical methods has generated greater interest in the field. Today, some of the most prominent scholarly and policy work in the field of civil conflict is driven by advanced quantitative analyses, including multiple regression and spatial statistics, many of which are reviewed by this chapter.

It is interesting to note that this transition from broad theory to statistical methods and mapping has reduced the significance attributed to individual decision-making in armed civil conflict. An underlying assumption of all statistical methods is that conflict is, to borrow the concepts of Michael Howard, less like an eschatological crusade, with one leader rousing people to sacred cause, and more like cataclysmic weather, with battles coming and going based on the environmental circumstances (Howard 1989). As a result, it is not surprising that our studies of conflict have increasingly incorporated the same techniques employed by ecologists to assess environmental risks, such as soil erosion, landslide risk, wildfire development, and other environmental risk (Ayalew and Yamagishi 2005; Gritzner et al. 2001; Ohlmacher and Davis 2003). These techniques – multiple logistic regression, GIS mapping, and spatial statistics – are increasingly common in civil war studies (Buhaug and Gates 2002; Buhaug et al. 2005; Buhaug and Rød 2006; Hegre and Raleigh 2005; Weidmann et al. 2006). Other techniques, such as cellular automata, which have already been used in modeling wildfires and simulating crime patterns, may prove useful in assessing the likely spread of civil conflict over time and space and thus assist in handling the complicated methodological and policy issue of assessing conflict contagion (Trunfino 2004; Anselin et al. 2000; Baller et al. 2001; Liang 2001; Buhaug and Gleditsch 2005). These techniques are not fully reviewed in my study, both because of the scope and limitations of the overall project and because they were not ultimately incorporated in the project's research model, but they could serve as fruitful avenues for future research.

My literature review in Chapter 2 separates previous work into either aggregate studies, which do not consider sub-national data, and disaggregate studies, which use a sub-national unit of analysis. As civil war literature has progressed, aggregate studies defined areas where findings have been robust, but the latest series of studies using disaggregated data began to question those findings. Disaggregate analysis included data more localized and specific data, thereby allowing researchers to question aggregate findings that natural resources are linked to civil war by testing specific natural resources, such as diamonds (Collier and Hoeffler 2004; Lujala et al. 2005). he course of these disaggregate studies, thus, leaves open questions about how findings that were theoretically and statistically robust when studied at a national level can have become so difficult to reproduce when studied at a sub-national level. This latest phase of research, which focuses on disaggregated country data, is far from complete, and this review highlights areas where further theoretical development and research data can hope to reconcile the emerging disputes. The following review is organized to provide a general overview of the transition in methods from theoretical and qualitative assessments, through country-year statistical regressions to disaggregated and geo-referenced studies. After the general overview, discussion proceeds by concept or variable in the following chapter, if the concept has been consistently operationalized, and starts with the theoretical basis for believing that the concept or variable could be significant and then walks through a few efforts to test the variable. This chapter does not attempt to address all that has been written on each of these concepts. Rather, the literature review aims to demonstrate how assessments have evolved and whether the evolution in prior methodology has further confirmed our understanding or created a need for additional research.

Having reviewed the major transitions in the past literature, I aim to use my thesis project to improve on the existing methodologies for studying violent civil conflict, to better understand the factors that determine when and where violent civil conflict occurs, and to assess whether the correlates of conflicts are distinct based on the aims of the parties to the conflict or the intensity of the fighting. The methodology for my thesis project is cross-national and empirical, and builds on major studies published in the past five years in this area (Fearon and Laitin 2003b; Collier and Hoeffler 2004; Buhaug and Rød 2006). The approach pursued by this project will contribute to the field by integrating aggregate political and economic information with disaggregate geographic information, incorporating intensity as a measure of conflict, and aiming to build an integrated model of state-level and local-level factors driving conflict. At the end of my thesis, regression results involving both state-level and local-level factors demonstrate how this process can be achieved. Previous studies have typically addressed civil wars as one group (Collier and Hoeffler 1998; Elbadawi and Sambanis 2002; Raleigh and Hegre 2009; Goldstone et al. 2010). Though some authors pull out "ethnic" wars or imperial-colonial wars (Sambanis 2001; Fearon and Laitin 2003b; Collier and Hoeffler 2004) from the broad category of "civil war," the literature reveals little distinction between different types of conflict or different levels of intensity. In addition to integrating time-sensitive political and economic data with location-sensitive geographic and environmental data, one of the central contributions that this project will make is to the field stems from the decision to separate conflicts by their type – either territorial (secession) or governmental (revolutions) – and their intensity – either low-intensity conflicts with less than 1000 battle-related deaths per year or civil wars with 1000 or more battle-related deaths per year. This separation allows us to identify factors that may significant for one type or intensity level but do not have broader significance. The study will be limited to Africa, which has seen the highest concentration of violent civil conflict in the past fifty years, and, perhaps consequently, inspired several comparable studies over the past ten years. The improved methodology should provide a framework for further research both beyond Africa and for examining other conflict patterns, and the findings from the study suggest policy responses for preventing violent civil conflict in Africa.

After reviewing the literature available for empirical studies of civil war, this project focused on research questions that appeared largely unaddressed or where progress could help to resolve outstanding debates in the literature. A number of discrepancies between the findings of aggregate and disaggregate studies, and those discrepancies may result in part from the different research approaches. As a result, the research questions that this study hopes to address are:

1. **Combine Aggregate and Disaggregate Data:** Is there a reasonable method for including data traditionally associated with aggregated country-year studies, such governance or economic data, in disaggregated studies, which have traditionally focused on geographic and environmental variables with no temporal variance?
2. **Differentiate Conflicts by Aims and Intensities:** What analytical insights are achieved by differentiating conflicts by their aims (territorial secession versus government capture) and their intensity (low-intensity conflict versus civil war)?
3. **Evaluate Specific Relationships for Variables:** By considering both aggregate and disaggregate variables using geo-referenced data, what relationships between political, economic, social and geographic variables and conflict emerge? Some of the major considerations include (but are not limited to):
    a. Does a pattern emerge that population density contributes to the risk of civil conflict in some areas and a deterrent in others?
    b. Does a pattern emerge that rough terrain (though either mountains or forests) are contributors to the risk of civil conflict in some areas and deterrents in others?

c. Does a pattern emerge that mineral deposits (both fuels, such as petroleum, and non-fuel, such as diamonds) are contributors to the risk of civil conflict in some areas and deterrents in others?
d. What role do borders between ethnic groups play in the likelihood that conflict will onset or occur?

4. **Consider Onset and Occurrence:** What role do variables relevant for conflict onset play in the continuation of civil conflict? Do conflict onset and conflict occurrence have similar or distinct causes?
5. **Determine Validity of Distinctions Between Conflict Types:** Is there a distinction between the correlates of different types of conflict for the onset and occurrence of hostilities?

Over the past year and a half, I have built a comprehensive dataset of approximately 80 variables and more than 7.5 million data points geo-referenced to time periods in Africa to examine the theories developed in this thesis. The construction of this dataset was a sizeable undertaking, not least because some of the existing data that I had expected to use proved to be too coarse or ill-defined. As a result, I gathered some of the data from other outside sources (such as for inequality data, rainfall, or elevation) or generated it on my own (such as formulation of the spatial and temporal lag variables). In the development of some of these variables – such as capital path-distance – I made a concentrated effort to link the specification of the variable as closely as possible to the theory that it represents. For example, if the distance between a capital and any given point is relevant because it represents the challenges for state forces in the capital to reach outlying territory, then the additional hurdles posed by mountains and rough terrain should be taken into account. Thus, using computer technology, the capital path-distance was calculated and used rather than just a straight “as the crow flies” distance. On the most critical variable – the conflict polygons that specify the dependent variable – I am grateful to Halvard Buhaug and Jan Rød for their willingness to share such carefully constructed data. The resulting dataset combines political, economic, demographic, social, geographic, and environmental variables at a sub-national level across the entire continent of Africa.

With this dataset developed, I applied standard logistic regressions, with clustered errors, to variables in this dataset in order to assess a wide range of hypotheses on the correlates of civil conflict. The resulting analysis indicates that different conflicts – governmental and territorial – and different levels of intensity are associated with different political, economic, demographic and environmental conditions. In other words, while many civil war studies have sought to produce the one regression that explains civil war, that approach may not be appropriate for civil conflict. Rather, it may be that explaining civil war requires multiple regressions to specify the different types and intensities of conflicts. In particular, my results suggest that the debate sometimes suggested between “greed” explanations for conflict onset, in which an insurgent group takes advantage of a weak state, and “grievance” explanations for conflict onset, in which an aggrieved group clashes with the state because of low opportunity cost, are both viable – but for different types of conflict. Governmental conflicts, in which the insurgent group focuses on capturing the government, are generally more aligned with “grievance” explanations, while territorial conflicts, in which the insurgent group focuses on capturing territory through secession, are generally more aligned with “greed” explanations. While the alignment of these broad theories to broad categorizations of conflicts is not exact, this kind of result does suggest that a single regression approach to explaining civil conflict may not be appropriate, and identifying several explanatory regressions for specific types and intensities of conflict may be more productive.

CHAPTER 2: REVIEW OF AGGREGATE AND DISAGGREGATE STUDIES OF VIOLENT CIVIL CONFLICT

This chapter provides a review of the major aggregate and disaggregate studies of violent civil conflict that have been conducted over the past five decades. The review is by no means exhaustive but is designed to highlight some of the major findings from aggregate studies, which focus on the country-state as the unit of analysis, and disaggregate studies, which focus on a sub-national unit, such as a grid cell, as the unit of analysis. Both approaches have had their advantages and drawbacks: aggregate studies have more thoroughly assessed political and economic factors, while disaggregate studies have more closely matched theories of civil conflict to sub-national data. Many of these studies have attempted to discern a single analysis for explaining civil conflict, sometimes excepting "ethnic" conflicts from the examples. Some studies, such as Buhaug and Rød's "Local Determinants of African Civil Wars," differentiate governmental and territorial conflicts, but none of the studies reviewed assess different types of conflict and different levels of intensity to show the distinct explanatory variables for both.

Country-Year Data and Large-N Studies

Overview and Basic Findings

The country-year method is one of the major strategies developed to more clearly, precisely, and demonstrably test theories that had been percolating in international relations and political science about the causes for the onset and occurrence of civil war. The country-year strategy is a basic approach to assessing the likelihood of conflict by gathering data for a set of countries over a set of years and then performing a multiple regression in which the conflict variable is the dependent variable and a range of economic, social, historical, and/or environmental factors are the independent variables. The method requires a great deal of data for these regressions to become robust enough to present credible findings. Singer and Small initiated the collection of the necessary voluminous amounts of data through the Correlates of War (COW) project in the 1960s at the University of Michigan in an attempt to gather together a list of variables relevant to assessing conflict in a rigorous and scientific manner. As Singer wrote in describing the beginning of the project,

> Rather than encouraging scholars to pick and choose cases according to their own recollections and biases, an accurate record of *all* cases of a given phenomenon was necessary to test hypothesized relationships with broad generalizability *(sic)* across time and space" (Singer and Diehl 1990).

As the basic method – multiple regression – for analyzing the civil war data in most country-year studies is not complicated the primary hurdle for effective execution is collecting appropriate and accurately measured data to include in the regression. (The development of increased computing power and data processing software also assisted in the implementation of empirical studies of civil war.) In this regard, the Correlates of War Project began to overcome this hurdle leading to the publication of early findings that addressed both international and civil conflict (Singer and Small 1982).

In the late 1990s, work by Paul Collier, Anke Hoeffler, James Fearon, and David Laitin began to narrow the focus to civil wars to create a targeted search for economic causes of civil conflict (Collier and Hoeffler 1998; Fearon and Laitin 1999). This work presaged the more comprehensive studies that the published several years later (Fearon and Laitin 2003b; Collier and Hoeffler 2004). Two works that have come to form the central pieces in the modern literature on civil war are Fearon and Laitin's "Ethnicity, Insurgency, and Civil War" and Collier and Hoeffler's "Greed and Grievance in Civil War" (Fearon and Laitin 2003b; Collier and Hoeffler 2004). These articles reach partially different conclusions, but, despite their distinct results, both studies employ the same basic methodology by taking data compiled on a country-year basis and then running regressions on the data to produce outcomes. Both studies also include large numbers of country-years. Fearon and Laitin code over 6600 country-year data points and use as many as 6360 country-year data points in one regression, covering civil war from 1945-1999 including anti-colonial wars (Fearon and Laitin 2003b). Collier and Hoeffler perform their analysis in five-year periods (1965-1969, 1970-1974 and so on until their final period, which is 1994-1999). Using this larger time period, Collier and Hoeffler have only 1288 country-year data points (Collier and Hoeffler 2004). And, as Fearon notes in a reaction to an earlier draft of their paper, grouping into five-year increments leads to the exclusion of up to one-third of the war onsets because of missing data on some aspect of the five-year period (Fearon 2005). Even with some shortcomings, the large datasets in both studies produce findings that cannot be easily discounted.

While Fearon and Laitin argue that it is state strength that holds back violence that would otherwise occur, Collier and Hoeffler argue that it is the availability of benefits from violence that causes it to arise. Fearon and Laitin argue that,

> Conditions that favor insurgency – in particular, state weakness marked by poverty, a large population, and instability – are better predictors of which

> countries are at risk for civil war than indicators of ethnic and religious diversity or measures of grievances such as economic inequality, lack of democracy or civil liberties, or state discrimination against minority religions or languages (Fearon and Laitin 2003b).

In contrast, Collier and Hoeffler focus more on the different motivations for civil war, and they conclude that,

> Most proxies for grievance were insignificant: inequality, political rights, ethnic polarization, and religious fractionalization. Only 'ethnic dominance' – one ethnic group being a majority – had adverse effects…Opportunity as an explanation of conflict risk is consistent with the economic interpretation of rebellion as greed-motivated (Collier and Hoeffler 2004).

Both studies do have broadly consistent findings in a number of other areas, though, such as that nations with large populations, low income per capital, high levels of oil exports are more vulnerable to conflict. In one notable additional finding, Fearon and Laitin also partially reject Huntington's culturally and ethnically focused argument by noting that the presence of Islam has little impact on the likelihood of civil war, and arguing that the large number of young males, especially with local schooling rates, present in some cultural groups is better explained as a social and structural factor than by a cultural root (Huntington 1993, 1996; Fearon and Laitin 2003b).

Criticisms of the Country-Year Regression Strategy

The country-year study approach has allowed a much larger amount of data to be brought to bear on questions of what variables tend to be associated with the onset of civil war than previous crude statistical methods that focused only on conflict zones without considering peaceful areas. It also identified a range of statistically robust relationships, but it is not without concerns. One key criticism of the country-year method is that the approach generalizes significant variation in a state and makes it appear static. For example, Fearon and Laitin generate their values for the percentage of mountainous terrain in a country. The authors openly admit that “This does not pick up other sorts of rough terrain that can be favorable to guerrillas such as swamps and jungle, and it takes no account of population distributions or food availability in relation to mountains; but it is the best we have been able to do” (Fearon and Laitin 2003b). As Fearon and Laitin recognize, it is very relevant whether an insurgency occurs in the mountainous area of a country or in a relatively flat area of an otherwise mountainous country. Meanwhile, Collier and Hoeffler, in their study, reference the importance of primary commodity exports, but do not include non-petroleum mineral resources, which is often cited as a source of financing for secessionist conflict. Even if Collier and Hoeffler had included mineral resources at a statewide basis and found a significant results, substantial debate would likely remain in the rest of the literature over the significance of primary commodity exports, either because of sensitivity to the procedure for imputing missing values (Elbadawi and Sambanis 2002) or because oil is still viewed as a underlying driver for the significance of primary commodity exports (Hegre 2002; Fearon 2005). As several authors have noted, the location of mineral resources is particularly important because, in contrast to oil or cash crops, mineral resources generally require comparatively less control of the state infrastructure to generate revenue, therefore making them a good source of funding for secessionist conflicts if the resources are distant from the capital (Le Billon 2001b; Ross 2004b). This unresolved issue – the impact of resource location on conflict – has been one of the areas where disaggregated studies have been most useful.

Using the country-year method has an impact on the overall direction of the analysis because it tends to lead to the selection of statewide variables to explain conflict because the aggregate approach cannot easily accommodate sub-national or geo-referenced variables. For instance, Fearon and Laitin use non-contiguous state, a binary measure for oil exporter status, new state status, instability (measured at the political center), form of government (also measured for the central government) as variables.[2] All of these variables – and per capita income when interpreted as a measure of central state strength – remain constant across the territory of the state, often measured as binary variables. (For example, a state either is or is not an oil exporter.) While there are regions within a country that might benefit more from oil exports than others, such as Scotland within the United Kingdom, the variable is taken to largely affect the development of the state. For instance, Le Billon argues that oil exporters generally have less sophisticated methods of raising revenue from their people and therefore are weaker overall (Fearon 2005; Le Billon 2001a). As already noted, Fearon and Laitin struggle to integrate variables that are not constant across the territory of a state. Fearon and Laitin do not use variables such as mineral resources, whose value to an insurgency may be related to their distance from the capital – proxying the ability of the state to control those resources. Using statewide variables fits within Fearon and Laitin's general framework that state-centric power determines the likelihood of civil war, and the absence of disaggregate variables makes it difficult to observe the potential impact of less state-centric variables.

Collier and Hoeffler's work focuses primarily on the intra-country dynamics that can lead to the development of rebel movements, such as unusually low opportunity costs

[2] "For oil exporters we marked country-years in which fuel exports exceeded one-third of export revenues, using World Bank data." (Fearon and Laitin 2003, p81)

for becoming a rebel. Collier and Hoeffler operationalize these concepts with variables such as male secondary schooling rates and the logged income per capita. While these statistics may be accurate for the nation as a whole, they are not necessarily good representations of the concepts that Collier and Hoeffler are seeking to understand. GDP per capita, for instance, varies significantly across the territory of a country, and it is definitely conceivable that low income statewide could still spark secessionist movements by the more economically advantaged, such as in the Somaliland portion of Somalia. As a result of relying on national-level to assess sub-national trends, the strength of Collier and Hoeffler's analysis is limited.

While most of these criticisms have focused on the representation and analysis of independent variables, such as income and rough terrain, there is also concern about the ability of the country-year method to adequately assess dependent variables regarding where conflicts occur. The country-year method struggles to adequately account for multiple conflicts within a single country. In the country-year method, there is no differentiation between widespread versus localized conflict, nor is there a method of handling simultaneous conflict in different locations or occurring within a short period of each other. As Collier and Hoeffler note, "If a war ended and another one started in the same [five year] period we coded these events as one" (Collier and Hoeffler 2004). This approach might be justifiable if multiple wars in the same period could be attributed the same prolonged conflict. Collier and Hoeffler's approach would be appropriate for the Sri Lankan civil war, a ceasefire ended most open conflict in 2004, but fighting between the same sides – the Liberation Tigers of Tamil Eelam and the Sri Lankan government – resumed in 2005. However, a number of cases exist within the Coellier and Hoeffler listing, when two separate and distinct civil wars occurred in a country within a five-year period. For example, in 1971, Pakistan fought a civil war that resulted in East Pakistan becoming an independent state, Bangladesh, and in 1973, Pakistan fought a second civil war when tribal leaders in Baluchistan attempted to secede. These conflicts are meaningfully distinct – occurring with different opposition groups and separated by thousands of miles in non-contiguous territories. While it could be argued that the secession of East Pakistan emboldened separatist Baluchis to rise up against Islamabad, there is little logic or value in combining two separate civil wars with different causes, circumstances, and key actors simply because they both fall in the same five-year period. Moreover, because Collier and Hoeffler's analysis focuses on the incentives for rebels, combining information about Bengalis in Bangladesh with Baluchis in western Pakistan in the same period obscures the distinctions between the two conflicts. Last, it is doubtful that the same environmental variables exist for the wet lowlands of Bangladesh and the drier areas of western Pakistan that are rich with oil and gas.

Ultimately, the aim of all civil war analysis is not just the production of advanced statistics but contribution to useful policy instruments. Analysis at the country-year level is useful for suggesting policy options that target the central government – often either making it more democratic or less dependent on resource rents. Aggregate analysis is also useful for those organizations, such as the World Bank (which sponsors much of the key research), that are making policy at a global level, where the intricacies of precisely where a conflict occurs become less immediately relevant than the broad strokes of which countries are at most risk. However, in many cases, the country-level analysis does not always produce particularly clear policy recommendations, particularly for large countries such as Russia, China, Brazil or India that often face address multiple potential civil conflict situations in different locations across their territories. Similarly, for the many organizations that seek to reduce civil war risk in just one country, it is useful to have a sense of *where* within a country civil conflict – and even what types of civil conflict – are likely to occur. In these areas, deeper sub-national analysis, achieved with the use of geographical or disaggregated studies, provides a clearer understanding of the local conditions that are associated with conflict and allows for organizations to respond in a more focused and directed manner.

Geographical and Disaggregated Analysis

Over time and in response to some of the criticisms that are noted in the discussion above, a range of studies have turned to the use of disaggregated data coupled with spatial statistics in order to gain a more detailed understanding of the local factors involved in civil conflicts. Instead of looking at the world as separated into discrete country-year events, these studies separate the world into much smaller territorial areas, either by the political boundaries of sub-national districts or by using grids to create unit cells for analysis. This approach permits studies to address more directly theories about the sub-national causes of civil conflict, such as a natural resource "curse" that implies that natural resources cause conflicts around them or that rough terrain increases the likelihood of conflict. With a disaggregated or geo-referenced study, it is possible to assess whether conflicts actually concentrate around natural resources or rough terrain within a country.

For example, Fearon and Laitin code the country of Bangladesh as being broadly ethnically and religiously homogenous and generally flat during the period of the Chittagong Hills conflict from the 1970s to the 1990s, which is accurate for the nation as a whole. However, the Chittagong Hill Tracts, where most of the conflict occurred, is the only significant area of ethnic and religious diversity and uneven terrain in Bangladesh. Thus, while Bangladesh appears to be an outlier when compared to Fearon and Laitin's general findings on civil war, a sub-national analysis would likely confirm that the conflict in Bangladesh conforms to the general concept presented by Fearon and Laitin (Haring-Smith 2007; Fearon and Laitin 2003b, 2003a).

Prior to the development of regression analysis based on large datasets of gridded geographical information found in this project, earlier studies focused on general statistical analysis of the basic geography of conflicts. For example, in 2002, Buhaug and Gates provided an analysis of the size of a conflict – how much territory it would cover – based on data on the location and spread of conflicts that had been collected (Gleditsch et al. 2001; Buhaug and Gates 2002). Contemporaneously, theoretical enquiries were building regarding where conflict should be located relative to specific geographical features. For instance, if rough terrain or natural resource deposits were relevant to conflict at an aggregate level, does that mean that conflict is close to these features or is the relationship more complex? These questions are best addressed through disaggregated and geographical studies where the geographic features in question, such as mountains or diamond deposits, can be mapped to specific locations in a country (Collier and Hoeffler 1998; Fearon and Laitin 1999; Le Billon 2001b; Malaquias 2001). These two activities led towards the gradual synthesis of these methods and theories into studies that attempted to test whether the evidence from country-year studies of implicitly disaggregated theories. Disaggregate approaches could be used to assess theories from aggregate theories, such as whether conflict should occur where rebels have the advantage of rough terrain or that conflict should occur where people are poorest and therefore have the least opportunity cost of enlistment with rebels.

To test these theories, datasets had to be developed that assessed not just which countries had experienced conflicts, but also where within countries conflicts had occurred. Among the first of efforts to develop the necessary data sets was a collaboration between researchers at the International Peace Research Institute (PRIO) at Norwegian University of Science and Technology (NTNU) and Uppsala University, called the Armed Conflicts Dataset (ACD). Introduced in 2002, the Uppsala/PRIO dataset coded a broader set of conflicts than those that had been identified by the Correlates of War (COW) Project, which restricted itself to only those cases that had 1,000 battle-deaths in a year (Gleditsch et al. 2002b). The Uppsala/PRIO dataset reduced its threshold to 25 battle-deaths per year, but, more importantly for the purposes of disaggregated analysis, it coded a value for the latitude and longitude for the “conflict center” of each conflict and a radius, defined as “the largest geographic extent of the conflict zone from the center point during the course of conflict” (Gleditsch et al. 2002a). The COW Project data had only identified the countries engaged in conflicts, never where within the country the conflict occurred. For many international relations scholars who used the COW Project data for international relations analysis, assessing a country as a single unit was acceptable (Collier and Hoeffler 1998; Hegre et al. 2001a; Collier et al. 2004), but the Uppsala/PRIO dataset offered those scholars working on internal conflicts a much more refined understanding of the conflicts that they were studying. Granted, the concept of a conflict as a circle – with a center point and a radius – was a crude basis for analysis, but the increased detail permitted some of the first studies of the geography of conflicts. In specific, Buhaug and Gates were able to develop models for the center point and total size of a conflict, which depended on the size of the country, the proximity to borders, the presence of natural resources, and the duration of the conflict (Buhaug and Gates 2002). While conflict circles were an improvement on the aggregate measures of conflict, scholars developed more refined measures of conflict geography to draw broader cross-national generalizations on civil conflict. Rød and Buhaug developed "conflict polygons" that traced the extent of a conflict in a more precise manner than circles, and then coded the polygons on to a grid, assessing any grid cell that contacted a polygon as having a conflict (Rød and Buhaug 2005). The creation of detailed, geo-referenced data on the size and nature of conflicts allowed for information to be rasterized[3] and compared to independent variables – such as roughness of terrain – through regression analysis.

There have been very few disaggregated studies of civil conflict based on this approach to date, but the results of the extant publications challenge the findings of prior country-year studies. The broadest disaggregated civil conflict study, Buhaug and Rød's "Local determinants of African civil wars, 1971-2001," used 3206 grid cells of100x100km to cover the continent of Africa (Buhaug and Rød 2006), aligning the grid so that every country is represented by at least one grid cell. Using conflict polygons based on the Uppsala/PRIO dataset, grid cells containing at least one conflict between 1970 and 2001 were identified. Buhaug and Rød classified these conflicts as either territorial or

[3] Rasterization is the process of taking a set of polygons on a map and converting them into a collection of grid cells that can be used spatial analysis.

governmental, following the designations provided by the Uppsala/PRIO dataset (Gleditsch et al. 2002b). According to the definitions, territorial conflict is a conflict where the opposition group becomes a separatist group, seeking control of a section of territory but not the government they are fighting. In governmental conflict, the opposition group seeks to control the government that they are opposing (Gleditsch et al. 2002a). Within the grid, 474 of 3206 cells contained "territorial conflict," and 1044 contained "governmental conflict." The decision to distinguish the two kinds of conflict was driven by an underlying assumption that the factors that drive these two types of conflict are different. While the empirical analysis by Buhaug and Rød ultimately supports this argument, the authors do not present an assessment of the empirical results from regressions include both governmental and territorial conflicts together.

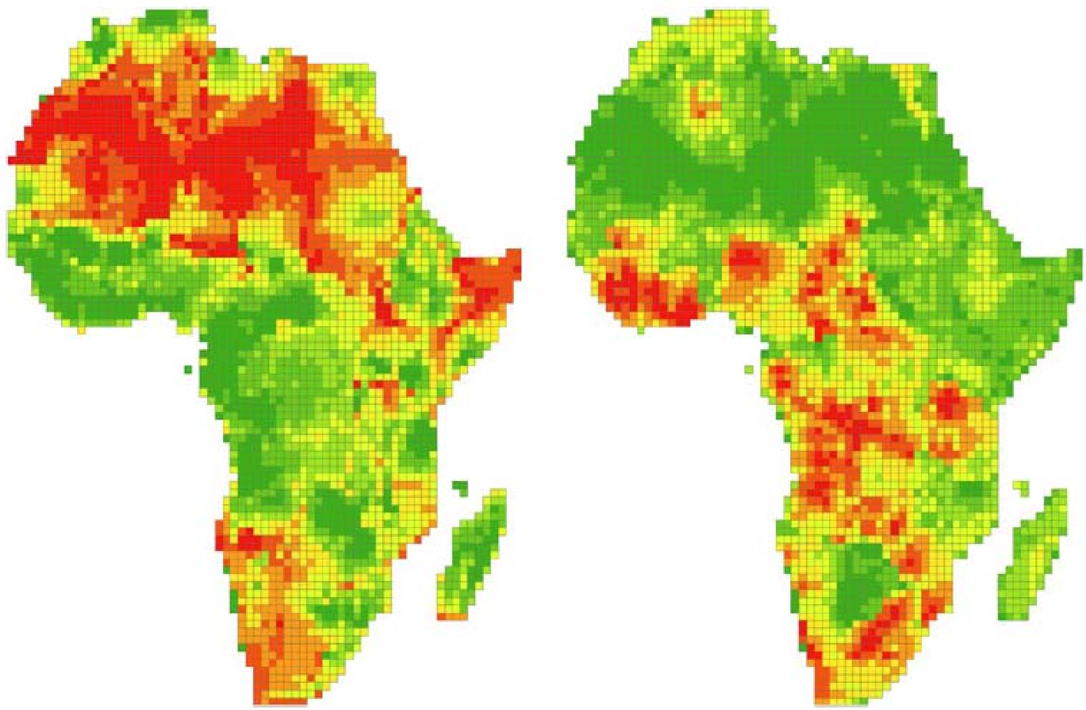


Fig. 2. Zones at risk of conflict. Note: The maps illustrate the relative risk of territorial (left) and governmental conflict on a scale from green (low risk) to red (high risk) as a function distance to the capital city, distance to the state border, population density, relative road density, share of mountainous and forested terrain, and distance to diamond fields. Darker shades signify higher risks. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Figure 1. Reproduced from Buhaug and Rød, 2006, figure 2 on page 331 of "Local Determinants of African Civil Wars" in *Political Geography* who constructed 100 km x 100 km grid cells for the whole of Africa. As marked in the reproduced caption, the maps show the results of their regression analysis for territorial and governmental conflict risk based regression results by grid cell.

Nonetheless, one of the major findings of the analysis is that the geographies of territorial and governmental conflicts are distinct. Government conflicts were located in cells that were "closer to the capital, further away from neighboring countries, more densely populated, have higher-than-average road coverage, contain more rough terrain, are closer to diamond fields and are less likely to contain a minority population" (Buhaug and Rød 2006). In contrast, territorial conflicts were "on average, closer to the border, further away from the capital, petroleum fields, and diamondiferous areas; they are less populated, less rugged, have less developed road network, and are more likely to contain a minority language than cells without territorial conflict" (Buhaug and Rød 2006). While Buhaug and Rød's empirical results were significant enough to believe that they might well be distinct, it is possible that there are some common underlying drivers of conflict that cause conflicts to start, and then, based on its location, the conflict takes on either a territorial or governmental nature based on secondary factors – such as proximity to the capital which would be a natural target and a valuable capture for any armed group that thought it was possible.

The mapping by Buhaug and Rød of the dependent variable – armed conflict – is done on a binary basis, but it is not clear from the Uppsala/PRIO dataset that it would be necessary to do so. There are two additional variables – duration and intensity – that could be used to generate a more gradated measure for conflict. The Uppsala/PRIO dataset marks how many years a given conflict lasts, a measure that could be used independently for analysis or combined in a measure with intensity to generate a "conflict severity" metric. For intensity of conflict, the Uppsala/PRIO dataset indicates three categories for conflicts:

1) "Minor armed conflict" - resulting in at least 25 battle-related deaths per year but fewer than 1,000 battle-deaths over the course of the conflict
2) "Intermediate armed conflict" - resulting in at least 25 battle-related deaths per year and at least 1,000 battle-related deaths over the course of the conflict but never more than 1,000 battle-related deaths in any single year, and
3) "War" - resulting in more than 1,000 battle-related deaths in a year of the conflict (Gleditsch et al. 2002b).

The Uppsala/PRIO dataset also provides the year-by-year aggregate total of battle-related deaths to justify the assessed intensity level of the conflict. While Uppsala/PRIO does not provide a measure of these intensities that would allow for distinctions between different areas of the same conflict, it might be worthwhile to differentiate conflicts by intensity to see what, if any, independent variables affect the duration or intensity of conflict. It also might help to address some of the problems of autocorrelation in which the presence of conflict in neighboring cells becomes the overwhelming driving factor in conflict in a given cell. Distinguishing by intensity would make it possible to make a distinction between two adjacent conflicts, one severe and one mild, which does not exist in the current coding of the conflict polygons onto the grid cells. In my project, grid cells are differentiated by intensity into low-intensity conflicts and civil wars, allowing for separate analyses based on intensity, but there is not a variable in which intensity is used a dependent linear variable. In previous studies, examples of using intensity as a dependent linear variable could not be found, possibly because of the difficulties of having sufficient data to specify intensity in a clear, linear fashion.

Other than finding a meaningful empirical distinction between territorial and governmental conflicts, Buhaug and Rød also highlight a number of results from the disaggregated study that diverge from the findings of country-year studies. Specifically, Buhaug and Rød find that territorial conflicts are more likely, in general, to occur in areas with less dense forest and more flat terrain, which runs contrary to a number of country-year studies (Fearon and Laitin 2003b; Hegre and Sambanis 2006). In contrast, governmental conflicts are more likely to develop near forests and rough terrain. Buhaug and Rød also find that proximity to diamond deposits does not increase the likelihood of territorial conflict – the opposite of what theory would suggest (Le Billon 2001b; Ross 2002) – but do correlate with a higher likelihood of governmental conflict. While the fact that more grid cells are coded for governmental conflict than territorial conflict might explain the split finding, it is still unexpected based on country-year models. Similarly, proximity to oil resources is not significant for either territorial or governmental conflict, even when non-oil producing states are excepted from the analysis, which runs contrary to one of the more robust findings of major country-year studies (Fearon and Laitin 2003b; Collier and Hoeffler 2004; Hegre and Sambanis 2006). While the presence of minority language – defined as a language that is not that spoken in the state capital – is a driver for secessionist conflict, it acts as an inhibitor for governmental conflict. As governmental conflict tends to be located near to capitals, the above-stated definition suggests a low likelihood of minority language areas overlapping with governmental conflict, so this result appears logical.

Criticisms of Disaggregated Studies

As Halvard Buhaug and Jan Ketil Rød note, shifting to a geographic analysis does not automatically address the problems created by the country-year approach (Buhaug and Rød 2006). Effectively, the shift to a geographically-driven analysis, in which the unit of analysis is a grid cell rather than a country, just reduces the size of the "countries" to much smaller units. But, there are still issues of the appropriate size, shape, and location for the units, since the size of the units can have an important impact on the results (Wrigley et al. 1997). While the borders of a country are politically defined and mostly unchangeable, much of disaggregated analysis relies on using grid cells that are defined by the researcher. Thus, the same discrepancies resulting from Fearon and Laitin's use country-year measures and Collier and Hoeffler's use of country-five-year periods can similarly arise when researchers use different sizes or shapes for the grid cells that they analyze. Broadly, this problem is referred to as the modified areal unit problem (MAUP), which consists of two major sub-components: scale effects and zonal effects. Scale effects occur when significantly different results are returned by analyzing data using larger grid cells with the same shape; for instance, shifting from square grid cells of 25x25km^2 to square grid cells of 50x50km^2. Zonal effects occur when grid cells of the same size are changed in their shape and produce significantly different results; for instance, shifting square grid cells that cover 625km^2 to hexagonal grid cells that cover 625km^2. The MAUP issue is one of the most significant problems with disaggregated studies that rely on grids, and spatial statisticians have yet to offer an easy solution. The simplest solution is to assess the data using a range of different sizes and shapes for the grid. While this solution works for the purpose of demonstrating the robustness of the results, it is also difficult and time-consuming. For example, in 2005, Buhaug and Rød presented the case for performing analysis at three resolutions (100x100km^2, 150x150km^2, and 200x200km^2) in order to address the MAUP issues raised by their study (Rød and Buhaug 2005), but it was not possible to implement this solution before the study was formally published as "Local determinants of African civil wars" (Buhaug and Rød 2006).

One potential method for working around this problem is to use defined political sub-districts of a country, as some studies do. While this technique may work for single country studies, it creates substantial data collection and data accuracy issues in cross-national studies. Also, in a cross-national study, the difference in the various sizes of sub-national districts has the potential to create significant issues. While the different sizes of countries affects country-year analysis, the impact of radical differences in the size of sub-national districts in a spatial statistical analysis would be considerably greater because location is a much greater factor in spatial statistical analyses than in country-year data regressions. Moreover, the internal sub-districts of individual countries are not stable historically, which would make it difficult to represent by district over a period of time for a geographic study.

One of the other major issues with much of the existing geographical analysis is autocorrelation. That is, conflicts tend to occur in neighboring locations, and it is not immediately clear whether this observation is because one conflict causes another or because both areas share similar traits that lead towards conflict. Most regressions performed on data in grid cells are not significantly different than those performed on country-years – the findings in each grid cell are taken as independent. If grid cells are not, in fact, independent because the presence of conflict in a neighboring cell leads to an increase in conflict in another cell, separate from any of the independent variables of that cell, then the analysis will become skewed. One method of addressing this issue is to include variables for spatial lag or neighborhood effects. For a spatial lag or neighborhood effect, an "independent" variable is included that measures the amount of conflict in the surrounding area to help differentiate the impact of nearby conflict from other variables being tested as correlates of conflict.

It is also important to note that disaggregated studies tend to do a less effective job than aggregate studies of incorporating time as a variable in the analysis. Both the Fearon and Laitin and the Collier and Hoeffler aggregate studies consider each country over a large number of years (Fearon and Laitin 2003b; Collier and Hoeffler 2004). So, for these country-year analyses, Sudan in 1970 is analyzed separately from Sudan in 1976. (While Collier and Hoeffler's decision to use five-year periods slightly decreases the consideration of the temporal effect, it is still strongly present.) For most geographical disaggregated studies, though, time is largely removed as a consideration. For example, in the previously discussed Buhaug and Rød analysis of African civil wars, all war zones from 1970-2001 are coded on the same grid (Buhaug and Rød 2006). The map displays one set of values per grid cell with no consideration of variation over time, or for the duration of wars, as we see with some other statistical analysis (Fearon 2004). There is also no proportionate weighting for areas that suffer perpetual conflict – such as Somalia or Afghanistan where almost all of the country-years are coded as war years. Instead, conflict in the Buhaug-Rød study becomes a binary condition, thereby ignoring the distinction between short wars, such as the brief civil conflicts in post-Soviet Eastern Europe, and the long, recurring conflicts in Somalia and others. While this lack of distinction can remove the often complicated issue of determining when exactly wars begin and end, it also precludes analysis of the factors contributing to conflict onset and resolution, which is particularly policy-relevant area of research. In essence, although disaggregated analysis gives us a much better sense of *where* conflicts occur, it struggles to provide a good account for *when* conflicts will occur. Broadly, this methodology strengthens our understanding of spatial effects in conflicts while weakening our understanding of temporal effects.

While it is certainly true that there is insufficient consideration of geographically varying variables in country-year studies, the reverse also holds some truth: there can be a tendency to provide insufficient consideration of nation-wide variables in disaggregated studies. In some cases, this stems from the difficulties of having little ability to incorporate temporally varying variables. With only one grid, there is no clear way to distinguish between Namibia before 1990, when it was part of South Africa, and Namibia post-independence. For 1990 and 1991, Namibia would be coded as a “new state,” which is generally associated with higher conflict risk (Fearon and Laitin 2003b). Nor is there an effective way to distinguish between apartheid South Africa, which was an anocracy and sparked a civil war between 1983 and 1994 (according to the Fearon-Laitin dataset), and post-apartheid South Africa, which is significantly more democratic and has not led to any civil wars. Similarly, other national level variables, such as Fearon and Laitin’s “stability” variable for any 3-point change in the Polity IV score for a country, are also difficult to include.

Syntheses of Aggregate and Disaggregate Approaches

While disaggregate studies made inroads to understanding the sub-national causes of civil conflict, the challenges associated with including data that changed over time, including key political, economic and social data, led to efforts to combine the aggregate and disaggregate approaches to civil conflict quantitative research. There have been some efforts to combine both state-wide and local variables but they are limited have little methodological consistency between them. Like disaggregate studies, information is plotted on a map and then rasterized to make assessments at the grid cell level. Unlike some disaggregate studies, though, these synthesized studies use smaller time periods – sometimes as short as a day – and then analyze the same territorial space at different points in time, incorporating updated social and economic information for each time period. To date, these synthesized aggregate and disaggregate studies have focused on one particular aspect of conflict – such as population effects – because of the time-consuming and complicated process of generating disaggregate data over multiple time periods.

In a worthwhile study on the impact of population on conflict, Hegre and Raleigh include both state-wide aggregate variables, such as the log population and log area of the country, and local variables, such the log population in a given grid cell as well as its distance to the state capital and international borders (Hegre and Raleigh 2005; Raleigh and Hegre 2009). The study did not include measures of regime type or data on overall country reliance on primary commodity exports, leaving room for further investigation of this technique of including both nation-wide and local variables. For the conflict data, Hegre and Raleigh drew on the Armed Conflict Location and Event Data (ACLED) dataset that tracked over 4000 battle events in 14 countries in central Africa over several decades (Raleigh and Hegre 2005). Rather than working with pre-set time periods, Hegre and Raleigh tested the "relative" risk of a battle event above the baseline risk of war at any point by comparing "snapshots" of the gridded data at the moment of any battle event. This "snapshot" approach generates a significant amount of data; Hegre and Raleigh included 100,000 grid cells measured during 2,530 battle events in their study. The authors elected to sample the battle events to reduce the computational hurdles of completing the analysis. The results of their study are useful in that they compare local and state population data in the same study, providing the useful insights that local population growth appears to reduce conflict and that overall state population is less significant than the local population. Hegre and Raleigh's research broadly indicated that conflicts were more likely in populated, accessible regions, although "conflicts are assumed to occur in faraway and inaccessible regions" – in part because of earlier aggregate studies.

Hegre and Raleigh's valuable synthesis of aggregate and disaggregate approaches has several challenges. First, the data is rasterized at such a refined level – 8.6x8.6km^2 grid cells (which often relies on information aggregated from 1x1km^2 cells) and with "snapshots" on the day of each conflict event – that volume of data could reach results could appear more significant than is truly warranted. Second, on a related note, Hegre and Raleigh do not appear to control for statewide variables – such as overall population size – not being sampled independently for each grid cell but only once per country. Last, the Hegre-Raleigh analysis admits that it is possible that the non-result for state population may occur from offsetting positive and negative trends. So, while the regression analysis appears to reject hypotheses of positive and negative relationships between state populations and conflict, it may be that <u>both</u> hypotheses are roughly equally true, offsetting each other into a non-result. Some of these problems – such as picking an appropriate grid cell size – are not easily resolved, but other concerns – such as controlling for the limited number of measurements of state population – can be addressed modifications in the statistical approach.

Another recent effort to assess conflict across both time and territory by Weidmann and Ward focused on Bosnia from 1992 to 1995 (Weidmann and Ward 2010). Their study used advanced statistical techniques to assess the specific districts in which conflict occurred based on correcting for both spatial and temporal lags with data collected on a monthly basis. The approach taken by Weidmann and Ward is quite interesting, but, regrettably, this research appeared too late in the process of my own thesis to test their statistical approaches. Nonetheless, their research could form the basis for future work on a larger dataset.

Both aggregate country-year and disaggregate geographical studies focus on other issues, such as oil production, but the operationalizing of the variables leads to different theoretical conclusions. For example, when oil export is operationalized as a binary variable, its primary theoretical implications are on the strength of the development of a sophisticated state capable of raising tax revenue (Snyder and Bhavnani 2005; Fearon 2005). When conflicts are measured for their proximity to oil resources, its primary theoretical implications are on sources of rebel financing or horizontal inequality (Brandsegg 2007; Buhaug and Rød 2006; Lujala 2005; Lujala et al. 2007; Lujala 2010). It may be possible to include both the binary data for some country-wide variables and the disaggregated data for local variables on a single grid for cross-national studies, so long as the areas being studied have not experienced significant changes in their international boundaries.

Where This Project Situates and Aims to Accomplish

This thesis aims to separate different types and intensity levels of conflict in order to reconcile results from an aggregate and disaggregate level. The findings of one set of disaggregated studies do <u>not</u> invalidate the results, or even necessarily the basic approach of country-year studies. Instead, the disagreements that emerge between disaggregated and country-year studies invite refinement of understandings of the situations observed. It is apparent from the existing literature that the same factors that lead to conflict in a country at a given time may not determine where in the country the conflict occurs. Although one might draw the conclusion that the factors that lead to conflict in a country would also determine where in the country it occurs, there is good reason to believe that it would not. While extreme conditions of poverty, rugged terrain, and large populations may weaken the central government of a state and thus create a political and physical environment for conflict (Fearon and Laitin 2003b), the basic needs and aims of rebels and peripheral government forces may push conflict into less extreme environments within a country. In essence, the process for conflict location can be considered a two-step procedure: one step when aggregate factors make a country prone to conflict and a second step that situates the conflict in the country; these two steps may not operate under the same tendencies or conditions. At the end of this project, an integrated model of state-level and local-level causes for conflict should suggest indications on where and when conflict is likely to occur, which may provide the basis for more targeted preventive action. Moreover, separating civil conflicts by their type and intensity helps to focus on what political, economic, geographic, or environmental factors may be most relevant to particular types of conflicts.

Conclusion

The progression from theoretical work on civil war to the sophisticated statistical approaches that are used in some of the most recent literature invites opportunities to try to harmonize some of the different approaches. As this study proceeds, it employs sub-national data to try to provide a more nuanced representation of some of the theoretical explanations for civil war. These explanations are inherently about intrastate relationships but have often been tested with only statewide data in previous cases. To ensure a closer match between the variables used and the theories tested, this project will use sub-national data in many cases.

In this project, data from both political and economic variables – traditionally the strong suits for aggregate studies – and geographic and environmental variables – traditionally the strong suits for disaggregate studies – will be used together. Combining these two streams of data allows for political, economic, geographic, and environmental explanations to be assessed side-by-side. By combining the data, as well as making other modifications to past approaches, the results from this project will not be completely comparable to past studies, but it will be more comparable to previous aggregate studies of civil war than other disaggregate studies because it separates out civil war from conflicts in general. (Buhaug-Rød group all conflicts with 25 or more battle-related deaths into one set, although they do differentiate by territorial and government conflicts.) While past studies have held largely to either a straight county-year format or a single grid cell measured over a long period of time, the Methodology section lays out a hybrid approach of using six five-year periods of grid cells that cover Africa over a three decade span.

CHAPTER 3: METHODOLOGY AND DATA SOURCES

Introduction

The aim of the methodology of this study will be to improve the existing disaggregated models of conflict in Africa by expanding the number of variables considered and integrating spatially varying disaggregate data with temporally varying aggregate data. In this section, I explore additional issues, such as more appropriate methods for specifying the conflict data and some independent variables. The methodology of this study does not aim to revolutionize the existing literature's understanding of civil wars, but it does explain how the framework for this project's disaggregated study will be more easily comparable to the existing cross-national studies that use a country-year methodology than previously published disaggregated studies. The work done to align the specification of the independent variables with the theoretical explanations that they represent required considerable technical work, but it is not designed to be the central contribution of this study to the existing literature. Rather, aligning the specification of independent variables with the theoretical arguments that they proxy should only bolster the broader primary work of this project. These refinements in the methodology and data specification contributes to the insights gained by differentiating conflicts based on their aims (territorial secession or government capture) and their level of intensity (low-intensity conflict and civil war).

In this Methodology section, I outline a "hybrid approach" that uses disaggregated grid cells taken in five-year periods over a three-decade span. This hybrid approach allows for both the political and economic data of aggregate studies to be included using the five-year periods while incorporating the disaggregate geographic and environmental data through the grid cells. There were no examples of this exact approach in the civil war literature that was reviewed, so the approach has been developed specially for this project. Hegre and Raleigh take a similar approach – integrating both spatial and temporal considerations – in their study of population concentration and civil war, using smaller grid cells and specific battles events as conflict data, but they do not study as broad a swath of territory and their unit of observation is the conflict event, not a defined time period (Raleigh and Hegre 2009). The hybrid approach taken by this study does partially resolve some of Hegre and Raleigh's criticisms of other approaches, including Buhaugh and Rød's study (Buhaug and Rød 2006), by allowing the area of the conflict to grow over time between the five-year periods rather than just measuring the total conflict area over the entire duration of the conflict. This hybrid approach also uses smaller grid cells (roughly 40km x 40 km) than Buhaug and Rød's study, addressing previously stated concerns that grid cells used are too large to be representative.

In addition, as a secondary endeavor, this study will briefly explore using Geographically Weighted Regression (GWR) as a method for evaluating local statistical relationships,[4] but the more detailed outcomes from that portion of the analysis are included in an annex as it does not play a central role in the work of the thesis.

After discussing the methodological approach, this section also lays out the individual data sources for each of my study. In general, the types of data used are similar to previous civil war studies, including economic, political, geographic, and environmental

[4] Initially, GWR was a more significant component of the thesis, but some of the early research and results from replicating other studies using GWR suggested that it would not necessarily be fruitful approach to follow in-depth.

data, but in many cases, I have altered the traditional specification of the variables or drawn from new sources of data. A full table summarizing these data sources is included at the end of the chapter.

Addressing Methodological Difficulties in the Existing Models of Conflict in Africa

As Buhaug and Rød acknowledge, there are several pitfalls with their methodology for disaggregated studies of conflict. This thesis will attempt to address two of them: 1) including data normally considered only in either aggregate or disaggregate studies by using several time periods, and 2) including consideration of intensity on conflict (Rød and Buhaug 2007). This section describes in detail the work that I did to make these two improvements in the basic disaggregate methodology.

Accounting for Temporal Considerations

As has been discussed earlier, disaggregated studies of conflict tend to take a single dataset that represents conflict events over a number of decades. Buhaug and Rød, for example, used the dataset to create a map of conflict over the period from 1970 – 2001. As a result of considering three and a half decades with a single dataset, the analysis focuses on geographical and other time-invariant considerations, to the detriment of considering some time-variant factors such as type of government and economic growth. While Buhaug and Rød may consider such time-specific variables as political system and economic growth important, their methodology precluded consideration of them. While it would be possible to go back and code "snapshots" of conflict for each year, effectively making a disaggregated country-year conflict study, this approach would be excessively time consuming and could still yield quite muddled results. Instead, my study uses data and approaches that already exist. Collier and Hoeffler divided their study into five-year periods, and collected data on that basis. It should be possible to use the existing data from Collier and Hoeffler, alongside the geographical data from Buhaug and Rød, if conflict polygons are assigned for each five-year period. This approach leaves the task of assigning the various elements of the conflict polygons that Buhaug and Rød created to five-year periods. While this approach required some new coding, it proved manageable. There are only 21 conflicts episodes that span one or more five-year periods, of the approximately 110 conflict episodes for internal armed conflict in Africa.

Table 1. This table presents the basic structural data on three key studies that preceded my thesis and the core data on my thesis.

Project	***Time Period(s)***	***Geographic Areas***	***Geographic Units***	***Time Units***	***Number of Conflict Episodes***	***Number of Total Observa tions***
Collier-Hoeffler[5]	1960 - 1999	Worldwide (161 countries)	Country	5 years	79 large civil conflicts	1,288 country-periods[6]

[5] (Collier and Hoeffler 2004)
[6] Due to unavailability of data, a maximum of 850 observations is used in any one regression in the Collier-Hoeffler study.

Fearon-Laitin[7]	1945 - 1999	Worldwide (161 countries)[8]	Country	1 year	127 conflicts (including 13 anti-colonial wars)	6,360 country-years
Buhaug-Rød[9]	1970 - 2001	Africa	100x100km grid cells	31 years	474 territorial conflict cells 1044 governmental conflict cells	3,206 grid cells
Haring-Smith Thesis	1970 - 1999	Africa	0.4x0.4 degree grid cells (roughly 40x40km)	5 years	1689 territorial hostilities onset cells 3046 governmental hostility onset cells	95,886 grid cell-periods

Thus, there are six datasets, effectively – one for each of the six five-year periods that are common between the Collier and Hoeffler study and the Buhaug and Rød study: 1970-1974, 1975-1979, 1980-1984, 1985-1989, 1990-1994, and 1995-1999. While it is unfortunate that data from 2000 to the present cannot be included, and the data will also exclude the first ten years (1960-1969) of the Collier and Hoeffler study and the last two years (2000-2001) of the Buhaug and Rød study. These years (1960-1969, 2000-2001) could not be used because the dataset required common years across the Fearon-Laitin, Collier-Hoeffler, and Buhaug-Rød datasets to allow for a common dataset to be completed.

[7] (Fearon and Laitin 2003b)
[8] Determined by which countries had a population of at least 500,000 in 1990.
[9] (Buhaug and Rød 2006)

For the growth data and other variables that refer to the prior period for 1970-1974 period, data from the 1965-1969 period in Collier and Hoeffler's study is used.

Including Intensity Variables

One of the other variables that is not considered by most disaggregate studies (and most aggregate studies) is the intensity of conflict. While it might initially seem daunting to assign a conflict intensity value for each conflict in each five year period, the Uppsala/PRIO Armed Conflict Database, version 4-2007, already includes information on conflict intensity in a conflict-year format. Though it is not practical to assign different intensities to different parts of the same conflict, which would be ideal, it is possible to assign different intensities to different conflict polygons in each five-year period based on the conflict-year data compiled by Uppsala/PRIO. While Uppsala/PRIO only distinguishes between "minor" conflicts which cause 25-999 battle-related deaths in a year and "wars" that cause 1000 or more battle-related deaths, this distinction is derived from the number of battle-related deaths that occurred in each conflict-year, which can be aggregated to assess some of the claims about intensity that have been made, such as by Weinstein, who has linked natural resources with more intense conflicts (Weinstein 2007). By aggregating the battle-deaths at each year, it is possible to

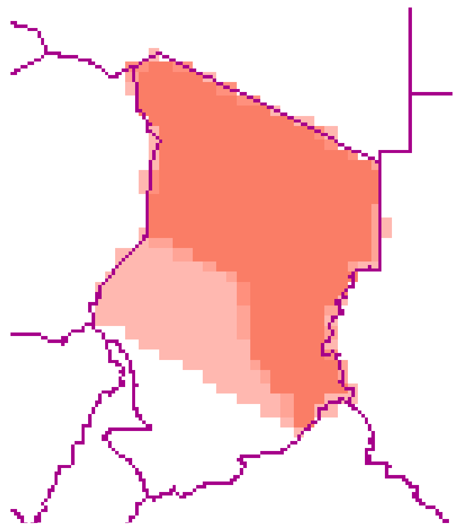


Figure 2. This pricture shows the differing levels of battle-related deaths in Chad from 1975-1979 as the conflict expanded southward in the country. Darker colors indicate more higher numbers of battle-related deaths for the conflict polygons.

generate different intensity scores in different parts of the conflict over the five year period and thus illustrate its growth. For instance, in Chad, the conflict that started in the northeast corner spread each year towards N'djamena, but it generated higher battle-death counts near its origin where the conflict has lasted longer and theoretically caused more deaths (see map of Chad battle-related deaths for 1975-1979). This reflection of the distribution of battle-deaths – with higher scores near the origin of the conflict – may not be completely accurate as the front-lines of the conflict shift, but it is a reasonable way of reflecting the available data. Moreover, a true concept of "intensity" has more to do with the number of battle-deaths per capita in a given area or per square kilometer for a given conflict. But, given the issues with assigning a weighted intensity across the conflict polygon, this study will make do with a simple score for the number of battle-related deaths from a conflict that included a given area. It is not an ideal measure, but it is the best estimation possible without data on individual battles, as is present in the Armed Conflict Location and Event Data (ACLED) dataset which is limited to 14 countries in central Africa (Raleigh and Hegre 2005). Due to the limited geographic scope of ACLED, though, the data was not used for my thesis.

After many different analytical approaches, the most reasonable approach, given the available data, seemed to be to classify the grid cells into different levels of intensity based on whether they qualify as "war" with 1000 or more battle-related deaths per year or "low-intensity conflict" with less than 1000 battle-related deaths per year. Although it would have been very interesting to be able to analyze the data on a more granular level, it would not have led to easy comparisons with the existing literature, which often restricts itself to studying war but not lower intensity conflicts. Moreover, this framework divided all conflicts into just two categories – war and low-intensity conflict – helps to resolve some of the questions about the quality of the data that has been gathered on battle-related deaths, where estimates for the annual deaths can vary widely. For instance, the Uppsala/PRIO dataset has very wide ranges between its low and high estimates of battle deaths for some years. In Ethiopia 1999, the dataset has a low estimate of 25 battle-related deaths and a high estimate of 999 deaths (Lacina and Gleditsch 2006). If my study were to focus on the precise number of battle-related deaths, this wide ranges would become a challenge, but there are fewer instances – only 30 out of 366 cases of countries in Africa between 1970-1999 where the low and high estimates for battle-related deaths span the dividing line of 1000 deaths per year (Lacina and Gleditsch 2006). In these 30 cases, I used the "best estimate" marked in the Uppsala/PRIO dataset as the number of battle-related deaths. Based on these considerations, consolidating my intensity measures into two categories – for civil war and for low-intensity civil conflict – was the most reasonable approach.

Methodology

Grid Unit Cells

The study relies on grid unit cells of 0.4 x 0.4-degrees to cover Africa. Buhaug and Rød established the basic principles of this approach; they used a grid of 100km x 100km cells to cover Africa, aligning the grid so that each country was represented by at least one grid cell (Buhaug and Rød 2006). In seeking to replicate this approach, though, it became clear that covering countries fairly with such large grid cells was quite difficult often resulting in less than intuitive representations, particularly in central Africa around Rwanda and Burundi, in western Africa around the Gulf of Guinea, and in representing unusual border patterns, such as Namibia's Caprivi Strip. While it originally appeared attractive to use the 100km x 100 km grid cell model, the smaller resolution allowed for less manipulation of the actual political map – for instance less crude representation of borders and less adjustment of capitals to fit the grid structure. While Buhaug and Rød do state that their conflict polygons are accurate at 50km, my choice to use slightly smaller grid cells alleviated a range of other inaccuracies that occurred when reviewing how approximately 50km x 50km cells would have fit along political boundaries with the economic and social data that is being included from Collier and Hoeffler's dataset, for which Buhaug and Rød did not have to account. (It is also worth noting that grid cells as small as 8.6 x 8.6km have been used in other studies when battle-specific information was available in previous studies (Hegre and Raleigh 2005), while country-specific studies have used cells as large as 76 x 76 km (Hegre et al. 2007).)

Defining the data based on longitudinal and latitudinal degrees allowed the conflict data to more easily interface with the external geographic and resource data. Unfortunately, the cells from these various sources are not identically sized. At their smallest – at the northern and southern edges of the continent – the grid cells measure about 35 km x 44.36 km, while on the Equator, the grid cells measure about 44.52 km x 44.36 km. This inequality is not a significant issue, though, since many variables represented by the cells, such as forest cover or mountainous terrain, are means and other variables, such as distance to the capital, are not affected by the overall size of the cell. These variables were averages from even more refined data so the size of the larger grid cell was not an issue. Moreover, the aggregate statewide data – such as economic growth or political system – that were mapped to every grid cell would not be affected by the slight changes in grid cell shape. Also, when conducting calculations to measure distances, the map was re-projected to the Eckert VI projection system, which is an equal area projection, so that measurements were made in consistent meter units rather than unequal degree units. Data was then projected back to the degree-based system for aggregation at the grid cell level. (Moreover, there is plenty of precedent for using unequal sized territories as observations – aggregate studies have shown China, Germany, and Malawi as equally weighted observations for years.)

To design this study appropriately, it is helpful to address the modified area unit problem (MAUP) (Wrigley et al. 1997), previously mentioned. Specifically, the research design should account for scale effects and zonal effects. The initial research article from Rød and Buhaug included only 100km x 100km grid cells because that was the only data generated at the time (Buhaug and Rød 2006). In their initial report, Rød and Buhaug randomly exclude 10% of their sample observations for cross-sample validation instead of constructing an entirely separate rasterization of the data (Buhaug and Rød 2006). In later work, Rød and Buhaug renewed their effort to tackle MAUP, and they made four different grids with cells of different sizes ($50km^2$, $100km^2$, $150km^2$, and $200km^2$), ensuring that in each grid at least one cell represents each country and relocating capitals as necessary to avoid having two capitals in the same cell. Rød and Buhaug also created a second $100km^2$ grid with a 52% overlap between grid cells (Rød and Buhaug 2007). I faced the same constraints as Rød and Buhuag in my research, and I have opted for the same work-around:

I reserved 10% of my observations for a cross-sample validation rather than re-gridding all of my data, which could not have been achieved on the available research timeline.

The interface between the shape data and rasterizing to a square grid provides for fairly easy conversion of data, and it makes the results more immediately comparable to other studies. At a technical level, the rasterization was completed in ArcGIS 9.1 with an ArcInfo license and then the rasterized data slides were saved as TIFF files and converted in to spreadsheets with ImageJ software using the Image IO plugin and exporting based on the XY coordinate of each grid cell.

Five-Year Periods

One of the great disadvantages of disaggregated studies is the difficulty of incorporating time as a variable. As described earlier, I constructed data to cover six periods for this project. At the aggregate and disaggregate level, data was organized into six periods, one for each five-year period between 1970-1999, integrating the data from Buhaug and Rød’s study of African armed conflict, Collier and Hoeffler’s study of civil war around the world, and other sources listed collectively at the end of this chapter (Buhaug and Rød 2006; Collier and Hoeffler 2004). In some cases, such as with the development of roads, it was not be possible to get accurate and distinct data for each of the six five-year periods. In these cases, I used the closest approximation , either by using data from the period directly prior or following, by imputing the missing data, or by using variables designed to proxy the same concepts or instruments. The decision whether to reuse data or impute missing values was made on a variable-by-variable basis, and the details of the data specification can be found in the following sections of this chapter. The six five-year periods are combined in a single dataset, and the grid cell level data are analyzed using a logistic regression with errors clustered at the country-period level. Using this approach, the results from the six five-year periods create a single regression analysis that reflects the overall trends of the data from 1970-1999.

Logit Regression with Errors Clustered at the Country-Period Level

I considered several approaches for analyzing the data in this project, and ultimately settled on a basic logit regression with errors clustered at the country level. While there were other approaches, such as multi-level modeling (described below), that were more appealing because they matched the basic data structure better, the logit regression approach is more consistent with previous literature, and it does work for the data that has been collected and for the variables involved. It is possible to use the same basic analytical techniques for all of the data involved; all of the data on the presence or absence of onset of conflict, occurrence of conflict, territorial or governmental aims, and intensity that is lower or higher can be measured as binary variables,. This consistency of approach has some appealing elements given that comparing results between different groups is a key part of understanding what, if any, differences there are between civil conflicts for territorial or governmental aims and those for low-intensity conflicts and civil wars.

Since this project works to assess data on both conflict onset – when does a conflict start – and conflict occurrence – when does the conflict exist, regardless of starting year – a modification on the basic methodology of the logit regression with clustered errors will be used for analysis of occurrence. For analysis of occurrence, a dynamic logit model will be used, also sometimes referred to as a Markov model. The approach with a dynamic logit model, which is also outlined and briefly used by Fearon and Laitin (Fearon and Laitin 2003a, 2003b), is to include an interaction term for each variable with the lag variable when assessing occurrence of conflict. The effect of these interaction variables is to distinguish the onset year from subsequent years of conflict in the regression. Ideally, this approach should produce a set of coefficients that are broadly consistent with the coefficients from the more limited logit regression that focused solely on onset years and with another set of coefficients that describe the continuance of the conflict past the onset year. In the dynamic regression results laid out by Fearon and Laitin, their results are consistent between the onset-only regression and the onset portion of the dynamic logit regression (Fearon and Laitin 2003a). Other scholars to use the dynamic model approach to analyzing conflict occurrence include Elbadawi and Sambanis (Elbadawi and Sambanis 2002; Theisen 2006; Vreeland 2008). Unfortunately, in my project, the impact of assessing both spatial and temporal lag variables – both of which tend to have strong significance – may impact the consistency of the results between the onset-only regressions and the onset portion of dynamic logit regressions. Nonetheless, the results from this method are likely to help to distinguish the impact of aggregate and disaggregate variables on conflict onset and conflict occurrence. And, to the extent that results for the onset-only clustered logit regression and the onset portion of the dynamic logit approach are consistent, it will help to bolster the robustness of those results.

Alternative: Mixed Multi-Level Regression

I also considered a second statistical method - mixed multi-level regression - for constructing models from the dataset, but opted not to use it because of some technical issues with the approach that could not be reasonably resolved. Initially, the mixed-multi-level approach was attractive because the data is being collected effectively at two levels – the grid-cell level and the state level – for each of the five-year periods, mixed multi-level regression appeared to satisfy the need for a statistical method that could accommodate the different collection levels. For instance, while there are over 95,000 thousand grid-cell separate data points for rainfall or elevation data over the six five-year periods, there are fewer than 350 separate data points for GDP growth. Those state-level variables are merely extended to the entire state's grid cells for that five-year period. As data at the state level has errors that are correlated across the space of the state, they need to be separated within the regression analysis. For example, all of the economic data in a country may be over-reported because of political biases within the finance ministry, so accounting for that common bias across the entire space of the country – all of its grid cells – is important. While multi-level modeling more explicitly takes the hierarchy of these biases into account, I opted to simply cluster the errors at the country-period level to account for these possible biases.

With mixed multilevel regression, the problem is similar to attempting to analyze data for student performance when a researcher has information about the individual students and about the schools that they attend. To differentiate the two levels, the data is grouped into separate steps within the same regression. For past studies on civil war in which information was gathered at both the local and state level, it is not clear that the regression process used adjusted for the less frequent collection of data for the statewide variables (Raleigh and Hegre 2009; Buhaug and Rød 2006). Hegre and Raleigh employed a "snapshot" approach in which the data is measured for every day that a relevant battle event occurs, which yielded more than 2,000 potential snapshots for analysis. Hegre and Raleigh then conducted a regression analysis to determine relative risk of conflict over a baseline risk level, but that method will not be employed in this study because the conflict data measures conflicts that stretch for years and is grouped in five year periods, rather than at the conflict-event level.

Unfortunately, mixed-multilevel regression could not be used for this project because classifying the approach into fixed and random effects could not be justified. From a review of the individual variables involved, a fixed effects approach was not appropriate because some the individual grid cells do not show any variation across time; some variables, such as those for elevation, level of forest density, or other geographic data, do not change significantly over time. Similarly, while there were likely omitted variables that exist at the country level – such as certain national policies or government leaders – but those variables are not fixed across the six time periods studied. Given the uncertainty of country and temporal effects in the dataset, clustering the errors at the country-period level in a logit regression was a more reasonable approach for the data, and one that is more consistent with past studies, such as Buhaug and Rød (Buhaug and Rød 2006).

Alternative: Geographically Weighted Regression (GWR)

Another approach that was considered, but not used for the final regression analysis, was Geographically Weighted Regression (GWR). As noted earlier, spatial autocorrelation has the potential to become such a significant factor in the regression results of disaggregated studies that it can overwhelm the significance of other explanatory variables. It has also been argued in the literature that the factors driving conflict in one area may not necessarily be the same elements that compel conflict in other areas. This situation is apparent both in the case of parabolic variables – such as primary commodity exports where a range around 30% appears to cause the greatest risk of war but is less significant both higher and lower – and in situations where the effect of a given factor can be different based on its geographic location, such as the proximity of diamond deposits to international borders. While some of these factors could be examined with an exhaustive list of interaction variables, GWR could be a more efficient strategy for observing these relationships, and it is explored in an annex of this study.

GWR, which is explored more thoroughly through replicating the Buhaug-Rød study, is not a new mathematical approach, but it has been previously employed to analyze conflict data. (The replication is presented in an annex because it did not become the main focus of my thesis.) GWR can reveal a number of potential connections between variables and their location that might not have been clear otherwise. Mixed multi-level regression allows for inputs for conflict from both the grid cell and the state level. At the very least, mapping the regression results from combining these factors at a disaggregate level will produce visual explanations and maps of the results that will be more intuitive and understandable for some key audiences, such as policy makers. (See Annex I for examples.)

Data Sources

In the original approach to collecting data for this thesis, my intent had been to gather data mostly from established sources, such as previously published studies on civil war. As the project progressed, though, it became clear that the sources of data would need to be expanded beyond the variables that had been considered by key prior work (Collier and Hoeffler 2004; Fearon and Laitin 2003b; Buhaug and Rød 2006). While I generally retained the economic variables from Collier and Hoeffler's study, the geographic variables have almost all been generated anew. I decided to regenerate the geographic variables because I needed the data at a finer resolution than provided by Buhaug and Rød's study. For instance, while Buhaug and Rød shifted the locations of some capitals because they fell in the same 100km x 100km grid cell (such as Cotonou in Benin and Lomé in Togo), the finer resolution of 0.4 x 0.4 degrees has allowed the calculations to proceed with no alterations to the original locations of cities. Similarly, some smaller countries needed to be designated a grid cell in the larger resolution, but in shifting to a finer resolution, no countries were so small that they were not automatically assigned at least one grid cell. (Comoros, the Seychelles, and Soe Tomé and Principe are the smallest countries, each with only one cell). A table that summarizes the number of cells in each country is below.

Country Name	Country Abbreviation	Grid Cells	Country Name	Country Abbreviation	Grid Cells
Algeria	ALG	1324	Madagascar	MAG	316
Angola	ANG	643	Malawi	MAW	63
Benin	BEN	63	Mali	MLI	675
Botswana	BOT	317	Mauritania	MAA	551
Burkina Faso	BFO	137	Mauritius	MAS	1
Burundi	BUI	14	Morocco	MOR	398
Cameroon	CAO	241	Mozambique	MZM	419
Central African Republic	CEN	321	Namibia (1990-1999)	NAM	455
Chad	CHA	673	Niger	NIR	637
Comoros	COM	1	Nigeria	NIG	462
Congo	CON	172	Rwanda	RWA	13
Djibouti	DJI	11	Sao Tome and Principe	STP	1
Egypt	EGY	575	Senegal	SEN	99
Equatorial Guinea	EQG	11	Seychelles	SEY	1
Eritrea (1991-1999)	ERI	63	Sierra Leone	SIE	34
Ethiopia (1970-1990)	ETH	643	Somalia	SOM	328
Ethiopia (1991-1999)	ETH	580	South Africa (1970-1990)	SAF	1159
Gabon	GAB	137	South Africa (1991-1999)	SAF	704
Gambia, The	GAM	7	Sudan	SUD	1298
Ghana	GHA	125	Swaziland	SWA	10
Guinea	GUI	133	Tanzania, United Republic of	TAZ	484
Guinea-Bissau	GNB	17	Togo	TOG	29
Ivory Coast	CDI	163	Tunisia	TUN	92

Kenya	KEN	295	Uganda	UGA	117
Lesotho	LES	19	Zaire / DR Congo	DRC	1190
Liberia	LBR	49	Zambia	ZAM	394
Libya	LIB	911	Zimbabwe	ZIM	208

In addition, using available public data on elevation, terrain and the locations of borders and capitals, it was possible to significantly improve on past measures of capital distance, roughness of terrain from forests and mountains by using GIS software. In prior studies, as is detailed below, expert judgments were sometimes used to identify terrain type, or approximations were made by averaging high and low points within a country, but satellite imagery and GIS software allows for each grid cell to have its own terrain measurements without averaging based on statewide data. Also using satellite imagery and advanced software, multiple specifications for distance were possible, including measures that took into account rough terrain along the path.

Figure 3. This map shows the rasterization of the countries for this project. Borders have not been included in this map in order to better show the impact of converting the linear borders to grid cell divisions.

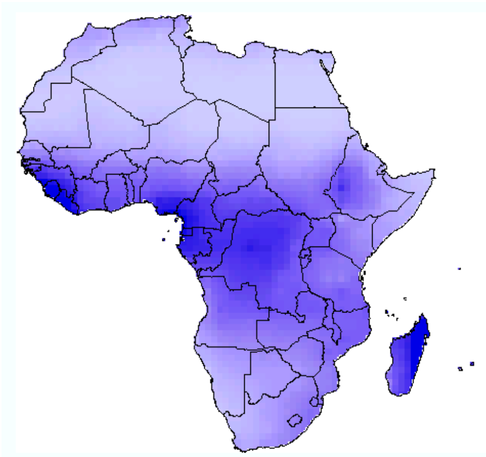
Economic Data

In my study, I used some economic data at an aggregate level because it reflects important overall concepts, such as state capacity, and because some variables, such as GDP per capita, have been consistently found to have a close relationship with violent civil conflict. I selected a range of economic data, such as GDP per capita and income inequality (Gini coefficient), that are drawn from the Collier-Hoeffler dataset for this study. For GDP, the data from the Collier-Hoeffler set varies sufficiently to provide useful information, but the income inequality ratio is only available for 27 of the 50 African countries that are included. Moreover, for the countries where income inequality data is available, it does not vary over the several decades measured by the dataset. As a result, it was necessary to look more broadly for data on income inequality.

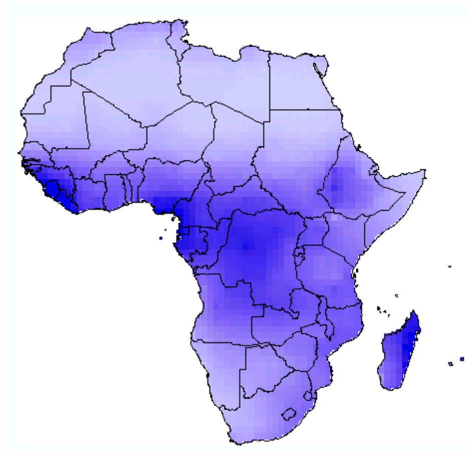
The United Nations University World Income Inequality Database catalogs income inequality with the available data for countries. While the available data for some countries, such as developed nations with strong economic monitoring agencies, is quite strong, the available data from most African nations is quite limited. For example, the UNU-WIDER includes only one measurement for Benin in 2003, compared to150 measurements, some gender-specific, for Australia from 1942-2001. Even this limited data, though, is an improvement on the available data in the Collier-Hoeffler set which has no income inequality information for Benin at all (*UNU-WIDER World Income Inequality Database, Version 2.0c* 2008). The other advantage of the UNU-WIDER dataset is that it provides a distinction between urban and rural income inequality in some cases. For example, UNU-WIDER provided different estimates for urban and rural income inequality in Botswana for 1986 and 1994, and similarly provided different estimates for urban and rural income inequality in Nigeria for 1975 and 1980. Of the African countries examined in my study, 15 states have Gini scores that are distinct for some sub-region in the UNU-WIDER dataset. In these few limited cases, economic data is available that is scored separately for urban and rural areas of a country in order to provide a slightly more disaggregate representation of the Gini factors for a given area. The data specific to urban and rural regions is not considered in my study for the purpose of regression analysis because it was not clear which territories were included in the “urban” and “rural” distinctions made by UNU-WIDER.

One of the major issues with economic data in studies of civil war is that it has an endogenous relationship with conflict. To help alleviate some of these issues, my study also includes rainfall data to provide a completely exogenous variable that is correlated with economic growth in Africa. The data for the rainfall is taken from the US National Weather Service 50-Year Reconstruction Project which reconstructed a dataset of monthly precipitation on a longitude/latitude grid with 2.5 x 2.5-degree cells (*50-Year Reconstruction* 2008). By interpolating the data to the 0.4 x 0.4-degree grid cells, it is possible to produce local rainfall estimates that serve as proxies for economic development. (Even using larger grid cells, such as Buhaug and Rød’s 100km x 100km cells, the coarser rainfall data grid would need to be interpolated to meet a finer resolution grid.) While this data is not likely to be a perfectly accurate reflection of economic growth because industrialized or urbanized areas will not have their growth as directly affected by rainfall, it should provide a more localized and exogenous measure of economic growth, as suggested by other studies (Miguel 2004; Jensen 2008). This data can be combined with local population data to generate local per capita rainfall figures. In this manner a potentially valuable dataset of local economic data can be generated that will cover the entire continent because the two data inputs – rainfall and population – are available continent-wide, unlike almost all of the economic and political data

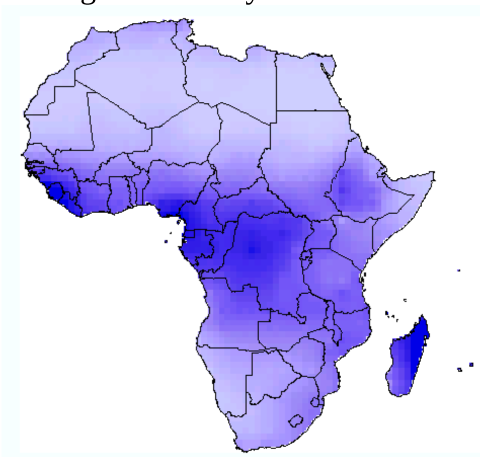
There are challenges in closely linking rainfall with local economic growth, as will be discussed again in the Analysis of Results. As there is little reliable local economic data in Africa on which to base conclusions, it is difficult to directly compare data on local economic growth with data on local rainfall, which can be easily gathered via satellites and technologies. While past scholars have taken rainfall as a local measure of economic growth, the varied economic performance across sub-Saharan Africa does not immediately appear to be directly indicative of rainfall patterns. Moreover, other scholars have raised concerns that there is not sufficient evidence to directly link rainfall and economic growth at the local level (Landau 1986; Bloom and Sachs 1998; Collier and Gunning 1999). At a minimum, though, local rainfall can likely be interpreted as a rough proxy for local resources since some amount of rainfall has been associated with important quality of life issues, such as proper nutrition and public health (Clover 2003).



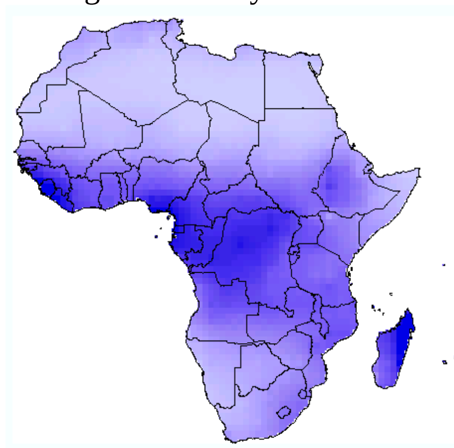
Average Rainfall by Grid Cell 1970-1974



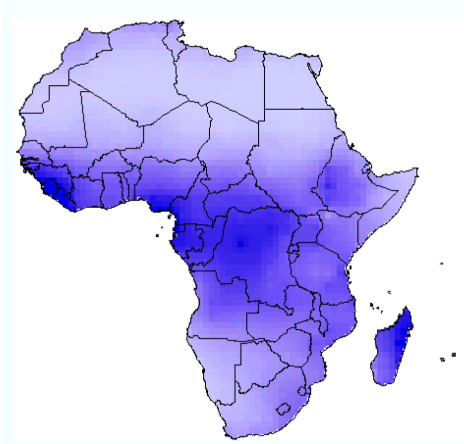
Average Rainfall by Grid Cell 1975-1979



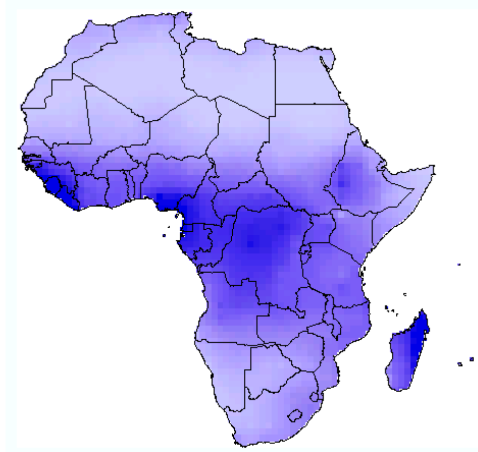
Average Rainfall by Grid Cell 1980-1984



Average Rainfall by Grid Cell 1985-1989



Average Rainfall by Grid Cell 1990-1994



Average Rainfall by Grid Cell 1995-1999

Figure 4. The set of rasterized data shows the rainfall data across the six periods covered. This data was taken from the US National Weather Service 50-Year Reconstruction Project and then imputed for values for grid cells.

Also, in an effort to broaden the number of variables considered and to check for robustness, this project also included data from the National Material Capabilities dataset produced by the Correlates of War project (Singer 1987). The National Material Capabilities data provides variables for energy consumption, iron and steel production, military expenditures, military personnel, as well as urban population and total population (making it easy to produce per capita measures). This information was not used in the Collier and Hoeffler, Fearon and Laitin, or Buhaug and Rød studies, although these authors do consider very similar variables. For instance, GDP and energy consumption are closely correlated. Given the many possible interpretations for the meaning of GDP – from interpreting it as a measure of state strength to a measure of the satisfaction of the population – it is helpful to use these additional variables to tease out which of the theoretical explanations appears to carry the greatest weight in explain the onset, occurrence, or intensity of conflict.

Political Data

For this project, a range of measures from the Polity IV dataset was used to provide political scores of regime type (Marshall and Jaggers 2007a). Polity IV (or its earlier predecessor Polity III) is the coding system used by both the Fearon and Laitin study and the Collier and Hoeffler study; Fearon and Laitin report trying different political scoring systems but note that Polity IV provided the best results when compared with conflict-year data on war onset (Fearon and Laitin 2003b). While transferring the scores is fairly straightforward, there is some contention over the proper way to process the scores. This contestation has been the topic of a considerable amount of literature already (Elkins 2000; Munck and Verkuilen 2002; Gates et al. 2006; Treier and Jackman 2008; Vreeland 2008; Goldstone et al. 2010). At the risk of adding to the already extensive discussion on appropriate handling of Polity IV scores, this project uses a slightly altered version of the Polity IV score (described below) to meet its analytical needs and to better match the score with the underlying theory.

Polity IV provides three ranges of scores that are considered autocracies, democracies and "anocracies" – a mixed regime type. It is conceivable that one could mark off those three categories and treat the political data as a categorical, rather than continuous, variable. Also, there are several kinds of discontinuities in the political regime type scoring – "foreign interruptions," "interregnum" or anarchy, and "transition." The issue of discontinuities came up in the literature previously when Fearon and Laitin took the suggestion from Polity IV to mark periods of interregnum or anarchy as 0 in their political system scores, effectively the same as a completely mixed regime type (Fearon and Laitin 2003b; Marshall and Jaggers 2007b). For Collier and Hoeffler, who used the earlier Polity III dataset (Jaggers and Gurr 1995), the creation of five-year periods of data also posed a difficulty. In their scoring, they averaged the scores for any available democracy score in the five-year period under study, leaving out any years that had a discontinuity. This approach has the advantage of not including the political scores that result from conflict when there are discontinuities, which happen often during conflict, but it has the disadvantage of occasionally tracking only the form of government that followed the conflict. For example, prior to 1975, Spain is scored as a 0 on the democracy scale and a 7 on the autocracy scale (for a composite score of -7). In 1975, the civil war began, and no score exists for democracy or autocracy, although the composite score ranges from -3 to 5 as it smoothes the transition to post-war democracy. In 1978, peace broke out, and Spain scored a 9 on the democracy scale and a 0 on the autocracy scale (for a composite score of 9). Collier and Hoeffler score the period from 1975-1979 as a 9 for democracy, but, in reality, that period included three years of bloody civil war when Spain could hardly have been considered a democracy. In cases when the entire five-year period was an interregnum or anarchy, Collier and Hoeffler simply scored the period as 0. This coding detail effectively makes anocratic scores (those around 0) endogenous with conflict, possibly explaining the result from Fearon and Laitin that suggests anocracy is associated with conflict (Fearon and Laitin 2003b).

A second set of issues surround the construction of the two major composite scores, democracy and autocracy. When summed, these composite scores comprise the final Polity IV score. Each composite score includes a range of different component variables – from scores for the regulation of the chief executive to the competitiveness of participation. While the composite scores have traditionally been the center of any analysis using Polity IV, it may be possible to use the component variables to narrow the lens on which components of “democracy” or “autocracy” create the conditions for conflict. If it is possible to assess likelihood for conflict through the component variables, then it could offer an insight into which of the myriad of conceptual explanations for the significant of political regime type to conflict risk is relevant. Alternatively, some scholars use sub-components of the Polity IV score – such as factionalism –to measure categorically with that consideration in mind (Goldstone et al. 2010). Both the overall Polity IV score and its sub-components face the same challenge some measures become endogenous with conflict because they include violence in their scoring. As noted by Vreeland, factionalism is one of those sub-components that include violence as a metric (Vreeland 2008). Scholars can assert that the measurements of violence in these subcomponents, such as factionalism, do not in fact matter because factionalism is still significant even when direct measures of violence are included in a regression analysis (Goldstone et al. 2010). Nonetheless, the issue of using a variable that partially measures violence on the right hand of the equation to estimate violence on the left hand of the equation is still not ideal.

Another strategy for including scoring for political systems is to compile sub-components to test for institutional consistency in the polity. Given a body of literature that suggests that institutionally consistent governments – i.e. with all sub-components indicating democracy or autocracy – are more stable than mixed regimes – i.e. some democratic and some autocratic sub-components, this approach fits with current understanding (Gurr 1974b; Muller and Weede 1990; Sanhueza 1999; Peceny et al. 2002). A challenge in taking that approach is that it ultimately relies on developing a composite index from merely a sub-set of sub-components for the Polity IV score, and some sub-components, such as selection of the executive, are not always linear in their coding and may be categorical.

Another of the concerns with the political data used is the issue of handling states who were not independent during some periods under consideration, which is a general challenge for studies in this period of time. For this purposes of my disaggregate study, the issue of independence is brought to bear by Angola and Mozambique. For instance, Angola and Mozambique were not independent of Portugal until 1975, and Namibia was attached to South Africa until 1990. Scoring the colonial relationship is not clear because no data from Polity IV is provided for the regime score of the colony. Collier and Hoeffler dealt with this issue by coding the colonial situations as 0, the midpoint between democracy and autocracy. Still, it does not seem reasonable to assume always a mixed system in the colonies. For example, the Collier-Hoeffler dataset marks Namibia in 1970 as a 0 (perfectly anocratic), but marks neighboring South Africa, which controlled Namibia, as a 4 (slightly democratic). The Collier-Hoeffler approach proves yet less appropriate when applied to intercontinental colonial relationships. To address this data coding issue with as much consistency as possible, I used the same approach that I used for the capital data – when questions were raised about whether to code Angola's capital as Lisbon in 1970 – will be used for the political data. In the case of Namibia, where the government is nearby, the same scores for South Africa will be used. In the case of European colonies, though, no data will be used since it is not possible to make an accurate assessment of the political situation in the colony – only in a far distant metropole.

Political data is one of the most difficult variables to justify simply averaging over a five-year-period, but as Collier and Hoeffler appear to have discovered, it is the most reasonable approach without moving to a country-single-year format. Thus, this project uses the time-series data constructed in Polity IV for overall regime type (ranging from -10 to 10), anocracy (the inverse distance of the Polity IV score to 0), and the various component scores that comprise the index scores. While this solution is not perfect, it will help to mitigate the endogeneity of reducing the democracy or autocracy score to 0 during periods of conflict.

My project also attempts a different, more unconventional handling of political data from Polity IV. The essential question that researchers ask in using political data is: "what political conditions led to conflict?" As a result, a second measure will use data that is generated by taking all of the scored years (i.e. not interregnum, anarchy, etc) prior to conflict in a five-year period and averaging only those scores. In the event that the first year of a five-year period is not scored, then the most recent scored data will be used. This system of coding generally leads to reasonable results – for instance, using the scores for 1977 Chad for the 1980 values because none of the years from 1978-1984 are scored. The most extreme case in this dataset is using 1991 Somalia for 1995-1999 Somalia and 1991 DRC for 1995-1999 DRC because there is no scoring for either country after 1991. This handling of the data means that there is no consistent relationship between the date of the political scoring and the year that the observation represents. Nonetheless, it does more closely present the information with which the theory is concerned – the political conditions prior to conflict.[10]

Proximity to Capitals and Borders

Regrettably, the proximity of each grid cell to the capital and to the border regions cannot be taken directly from the Buhaug-Rød dataset, as originally considered, because the borders and capitals of countries have changed over the period in question but only one capital distance is recorded in the Buhaug-Rød dataset, which covers all of the period from 1970-2001 with one variable. The instances that raise a coding question are the creation of new states, shifting of borders, and the shifting of capital locations. This did not prove to be a substantive issue for this study, as relatively few new states emerged in Africa post-1970.

[10] As noted later in the results section, this second scoring approach proves to be more significant to the results.

For the cases in which a new state did emerge during a five-year period, such as Namibia in 1990, then the area of the new state is coded using the distance from its former capital until the next period, unless if falls on a break year. In the case of Namibia, the 1985-1989 dataset reflects its distance from Pretoria, South Africa, but the 1990-1995 dataset displays its distance from Windhoek, Namibia. (Accordingly, the distance to an international border drastically decreases for the land that had been located between Pretoria and Windhoek, but, with the creation of Namibia, was now very close to the Namibian-South African border.) In the case of former European colonies, the distance from the capital will be recorded as the projected distance to the colonial capital. In the case of Guinea-Bissau prior to independence in 1974, the distance between it and Lisbon, Portugal will not be used; instead, it will be the distance to Bissau, since that represents the local point of greatest state strength for the government.

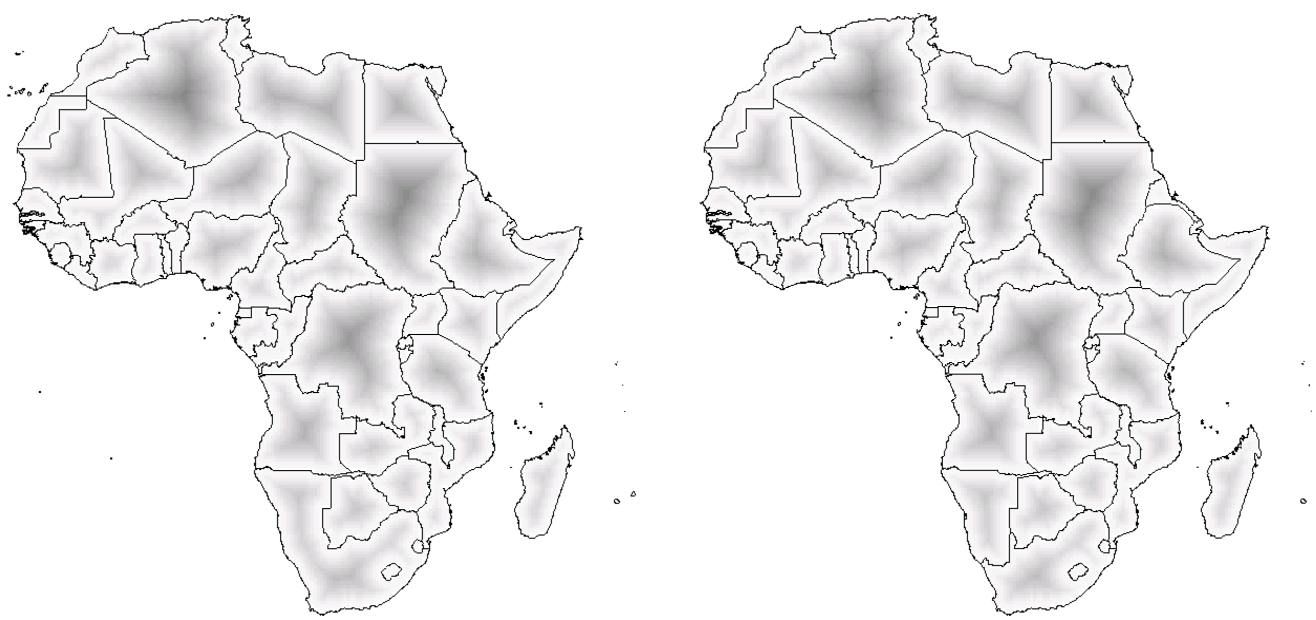


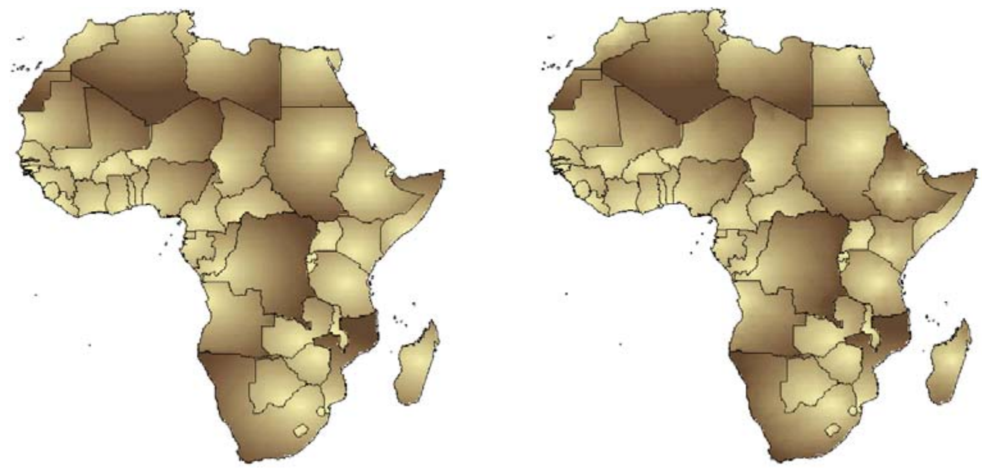
Figure 5. The maps above show the border distances for 1970-1989 on the left and 1990-1999 on the right. Note the changes in the borders for South Africa/Namibia and Ethiopia/Eritrea.

The shifting of capitals is a more intriguing situation because it reflects some of the underlying political dynamics in the countries observed. In a number of cases, the capital for a state was shifted since 1970. Côte d'Ivoire, for instance, moved its capital in 1983 from Yamassoukro, located in the southeast corner of the country to Abidjan, which is more centrally located. Given that Côte d'Ivoire suffered a civil war that was split largely between north and south, moving the capital north and closer to the center of the country has a meaningful impact on rebel's distance to the capital (and vice-versa). Similarly, Nigeria shifted its capital from Lagos, in the southwest corner of the country to Abuja in 1991. With nearly thirty thousand battle deaths occurring in 1970 from a conflict with the self-proclaimed Republic of Biafra in the southeast of Nigeria, the eastward shift of the capital may affect the propensity for conflict in the north and east. (Many thousand more deaths occurred from the Biafran rebellion in the 1960s, but that falls outside of the scope of this study.)

An important variation on the traditional method for measuring the distance to the capital is also added to this study. When assessing the distance from a capital to a given area, the theoretical explanation for its significance is often that distance makes it more difficult for the government to penetrate these areas, and therefore easier for rebels to seek refuge and build a social and cultural distinction between the center and the periphery. Given these explanations, it seems that a much more relevant consideration would be how difficult it is to reach a given area from the capital. In an analysis called a cost-path, distance is measured based on how difficult it is to reach the capital from a given area, taking into account elevation. This measure is a more appropriate specification for the underlying theory. It is important to note that the method does not just proxy the elevation, slope, or terrain for the area examined. It measures the accumulated distance of the easiest path from the capital to a given area, taking into account the elevation changes in between. In measuring the cost path distance from the capital to a given area, it is possible that both the capital and the area are perfectly flat, but there could be a substantial distinction between the straight-line distance and the cost-path distance because of a large mountain range that separates the capital from the peripheral area examined. Thus, the cost path distance assesses the true distance between the capital and a given area.

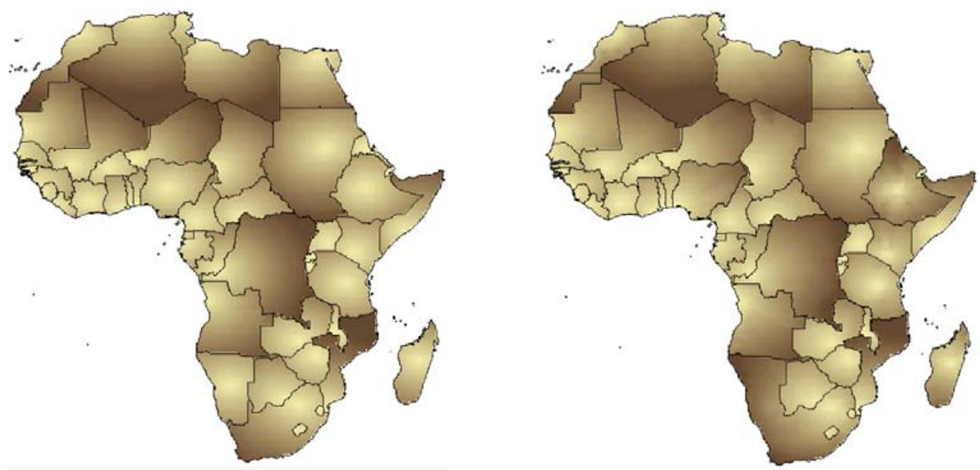
1970-

1979



1980-

1989



1990-

1999

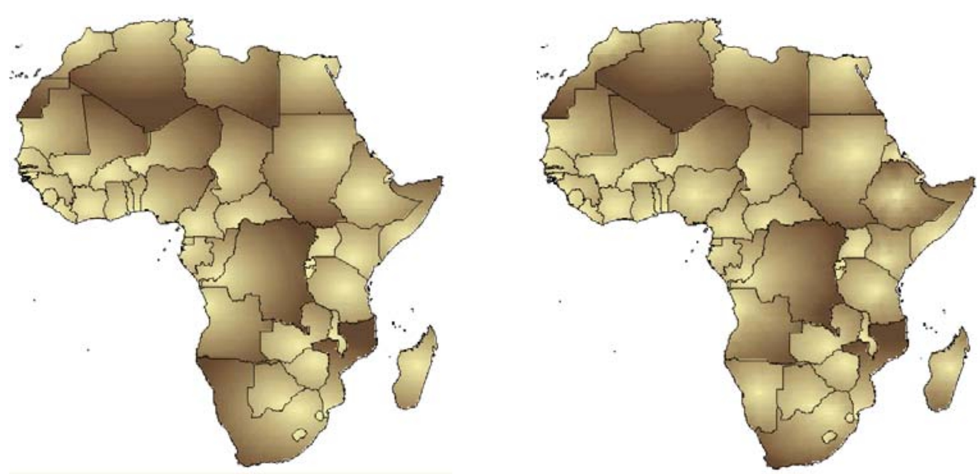


Figure 6. The maps above show the capital straight-line distance and the capital path-distance for the periods covered. Note the changes in capital location in Nigeria, South Africa/Namibia, Ethiopia/Eritrea, and Ivory Coast. Also note the impact of the elevation changes on capital distance in Ethiopia, Morocco, and Kenya in particular.

There are, of course, flaws in this system of measuring the distance from a capital to an outlying area. For instance, while the Sahara is flat, the intense conditions of the desert make travel particularly difficult. Similarly, the jungles of central Africa can add a degree of difficulty to travel through the region that exceeds simple elevation changes. It is tempting to construct a single metric, taking into account all manner of terrains and elevation changes, but given the somewhat arbitrary manner in which difficulty costs for various terrains would need to be assigned, the cost path distance for this study will simply account for elevation changes. This measure is not a perfect proxy, but improves upon the simple straight line distance. Future studies should explore the feasibility of using the financial cost of building a kilometer of road in any given area as a unifying metric.

Ideally, road networks would also be taken into account, allowing for a much more accurate assessment of the difficulty of penetrating into different parts of a state's territory. Unfortunately, no reliable digitized map of the roads for the continent of Africa over the three decades being examined exists. While it might be possible to use the map of roads from the *ESRI Digital Chart of the World* that Buhaug and Rød utilize, that map only measures the road network during one year – a snapshot that does not take into account key changes that may occur over the course of a decade or more.[11] While Buhaug and Rød argue that using a single road map for a study of several decades is acceptable because the relative ranking of the road density in different parts of Africa did not change significantly over time, using other local economic data appeared to be a more reasonable approach for this study. (It is worth noting that the reverse decision was made regarding ethnic data.

[11] For a prime example of the importance of planned transportation in changing the political and social dynamics of a potentially rebellious region, consider Tibet – an area located high in the mountains and very far from the capital of Beijing – and the planned railroad from Beijing to Lhasa that has dramatically reduced travel time to the outlying area and increased the number of Han Chinese who are able to reach the region.

Even though only one outdated ethnic map was available, it has been included in the study because there were no other viable, updated data sources on local ethnic composition.)

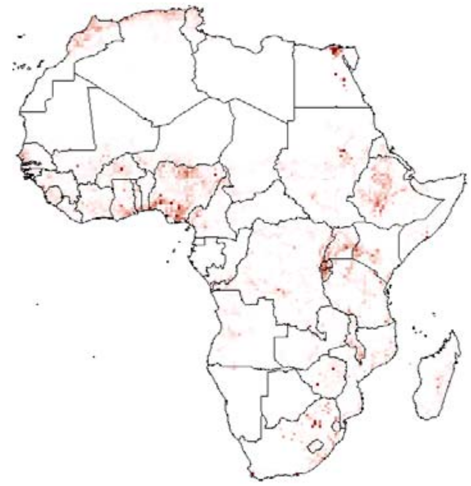
Poverty and Development

Researchers in both aggregate and disaggregate studies have struggled to identify appropriate variables to proxy the complex and important concepts of poverty and development. To proxy economic development and state infrastructure, Buhaug and Rød used road density data taken from ESRI's *Digital Chart of the World* ("Digital Chart of the World" 2007). This data was only present for a single year – 1997 – in their study period for African wars, from 1970-2001, but Buhaug and Rød cite information that relative road density as a rank order among countries did not change significantly over the three decades of their study in Africa (Herbst 2000; Buhaug and Rød 2006). The argument that the rank order of countries based on road density remains constant is a strong argument that using a single data source would not interfere with an aggregate, country-year study. However, it is less clear that the distribution of roads within a country would remain constant, creating a challenge for disaggregate, sub-national studies. One of the rebel arguments for secessionist conflicts, such as the Tuaregs in Mali, has been a lack of infrastructure investment in remote areas. Since it is possible the internal distribution of road networks experiences relevant changes over time, my study will not incorporate that data since insufficient sources are available to measure the changes over time. Rather than use the road networks data, though, it possible to rely on the economic development that is used in the Collier and Hoeffler study (Collier and Hoeffler 2004). Thus, rather than using a geo-referenced dataset of road networks as a proxy for economic development, it should be possible to use actual economic data instead. And, since Buhaug and Rød only assert that the rank order between countries remains stable, transitioning to economic variables that are country-wide should not change the level of precision too greatly.

Population

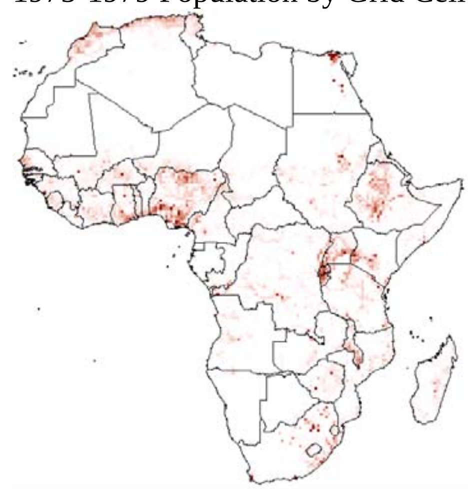
Conflicts are waged between groups of people, so the size and concentration of population is an important factor to consider in modeling conflict. The existing Buhaug-Rød model for African civil conflict uses population data from the United Nations Environment Program’s Global Resource Information Dataset (UNEP-GRID). The UNEP-GRID provides geo-referenced data on population density for each decade from 1960-2000 for Africa and Latin America, and for 1995 for Asia ("Global Resource Information Database" 2007). There is also a global population density grid provided for 1990. The UNEP-GRID Dataset, which comes in the form of a grid with cells that are roughly 0.04 degrees square also serves as the source of data for my study.

1970-1974 Population by Grid Cell

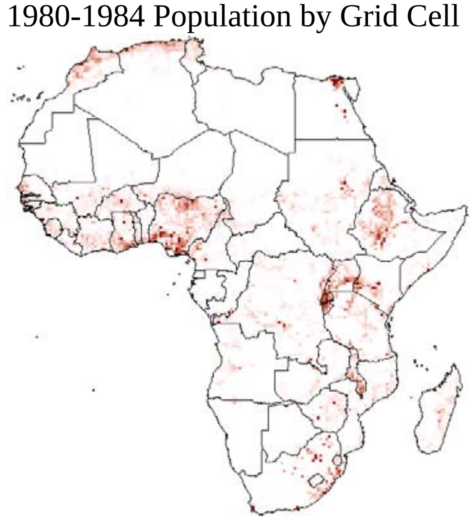


1975-1979 Population by Grid Cell

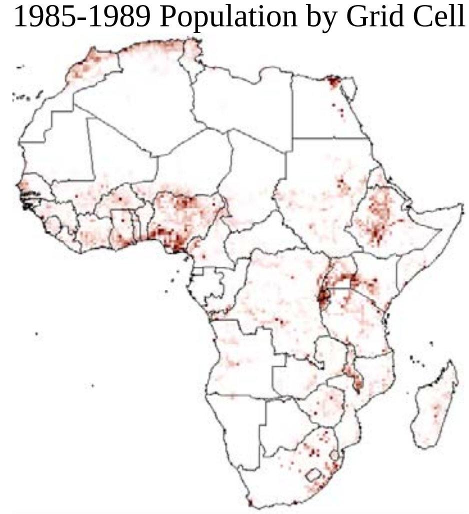
1980-1984 Population by Grid Cell



1985-1989 Population by Grid Cell



1990-1994 Population by Grid Cell



1995-1999 Population by Grid Cell

Figure 7. The maps above show the population by grid cell for the five-year periods used in my thesis. The scale for the variables has been changed so the at the darkest cell is set at 10,000 people / grid cell, even though the maximum ranges from approximately 35,000 in 1970-1974 to over 80,000 in 1995-1999.

The data used by Buhaug and Rød were the sum population in each grid cell for all of the decades in their study. When the data is broken down into five-year periods, it is not necessary to aggregate the population data in this manner over a longer period of time. Instead, the each of the five-year periods reflects a much closer representation to the population distribution at the time that conflicts occurred interpolating from each decade-level data to fill in the half-decade intervals that are missing for the five-year periods of this study. The method used for interpolating assumed exponential population growth (or decay in the handful of situations where population decreased over of the course of the intervening decade). In order to perform this calculation, it was necessary to add 1 person to every area, eliminating the possibility of infinite growth in the event that an unpopulated area became populated. This adjustment also eliminated the possibility of infinite rainfall per capita and other calculations based on the local population. After interpolating the mid-decade population figures, the data was then aggregated to the 0.4 x 0.4-degree resolution by taking the mean of the cells in the finer resolution grid.

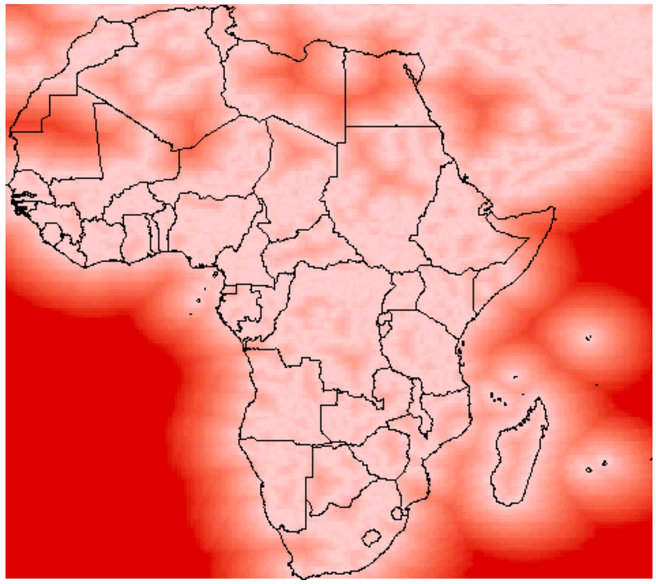
In addition to considering the local population, another measure of state population for my study was taken from the National Material Capabilities dataset at the Correlates of War project (Singer 1987). I chose to include both local and state population sizes is that a) local population in nearly fixed grid cells essentially amounts to local population density and b) there are different theories about why conflict would be affected by state and local populations. There is evidence at both the aggregate and disaggregate level that the overall population of a country increases the likelihood that civil conflict will appear in some part of that country's territory, either because the large population is too difficult for the state to manage or because the large population increases the chance of disaffected groups forming

(Fearon and Laitin 2003b; Hegre and Sambanis 2006; Buhaug et al. 2009). These basic structural arguments suggest that population patterns should be considered simply because of the increased population. Meanwhile, there are separate explanations for why population density or local population might affect conflict risk by changing the political dynamic or military balance in a country. For instance, in sparsely populated locations, there may be a less developed infrastructure, which could support either the greed argument that it provides safe havens for insurgents or the grievance argument that the lack of state investment promotes resentment and rebellion. A number of case studies suggest that both theories may be correct. For example, the Tuaregs in northern Mali and Niger, consistently cite the lack of investment in their regions and utilize the military advantage that local knowledge of the Sahara has given them to resist government forces (Keita 1998) Living far from the population centers – the capitals – of Mali and Niger, the Tuaregs do not fit the mold of a group who experienced conflict solely because they were one of many groups in a populous state, but rather the case study indicates that even in cases of *low* population, the political or military dynamic can develop to create conditions for conflict. Thus, in reviewing the role of population in conflict risk, it is possible that there are dynamics that would create conflict risk for both high and low population concentrations.

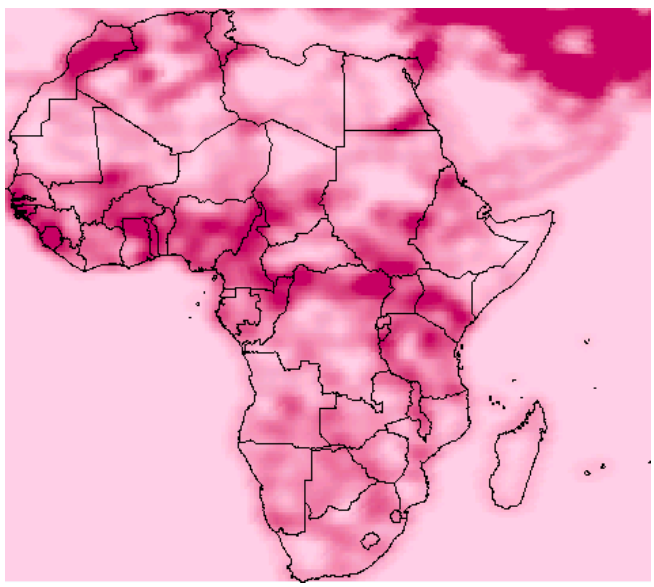
Ethnic Groups

Measuring ethnic groups with quantitative data is a fraught process; indeed, examinations of what constitutes an “ethnicity” or an “ethnic group” form the bases for other disciplines of study. Unlike more objective measures, such as elevation, determining how big an ethnic group is or where it is located is challenging at both a theoretical and practical level. Moreover, even if one could create a “perfect” quantitative measure of different ethnic groups’ size and location, natural movement patterns could change that arrangement relatively quickly. Still, ethnicity has been frequently cited in literature as a cause for violent civil conflict, so some effort must be made to capture the salient elements with quantitative data (Fearon and Laitin 1999; Sambanis 2001; Toft 2003; Buhaug et al. 2006). In this pursuit, data on ethnic groups and ethnic fractionalization have been measured in aggregate or disaggregate approaches.

Data on distance to an ethnic configuration border, not counting coasts as borders. Data taken from Geo-Referencing Ethnic Groups (GREG) program and extent of data extends beyond borders to show potential impact of coasts.



Data on kernel density of ethnic configuration borders, not counting coasts as borders. Data taken from Geo-Referencing Ethnic Groups (GREG) program and extent of data extends beyond borders to show potential impact of coasts.



In the aggregate approach, data on the percentage of each ethnic group as a composition of the whole state's population can be used to build a fractionalization index. This approach has generated a number of different measures for "ethnic fractionalization," but one of the most common is Philip Roeder's Ethnolinguistic Fractionalization Index (Roeder 2001). As the name suggests, the measure relies heavily on linguistic differences to define different groups. Roeder calculates the ethnolinguistic fractionalization of the country twice, once for 1961 and once for 1985, based on revised population data. Other key authors, such as Fearon and Laitin, have developed their own ethnic fractionalization indices using source data from *Atlas Naradov Mira* (described in detail below), the CIA Factbook, Encyclopedia Britannica, or US Library of Congress Country Studies (Fearon and Laitin 2003b). Fearon and Laitin calculated the ethnic fractionalization variable as "the probability that two randomly drawn individuals are from different ethnolinguistic groups" (Fearon and Laitin 2003b). Collier and Hoeffler also created their own measure for ethnicity, and they chose to specify the binary variable as cases when the largest ethnic group has 45-90% of the population, "a small majority," that Collier asserts makes exploitation by the majority more attractive (Collier and Hoeffler 2004). These approaches hinge on a state-wide scope to the analysis, since the aggregate measures of ethnicity do not present information about specific areas of vulnerability within a state's territory.

As aggregate estimates are based on more broadly reported population data, they are easier to revise and update than geo-referenced disaggregate results. Still, for the purposes of civil conflict analysis, aggregate estimates have several shortcomings. Most notably, aggregate studies do not reflect the arrangement of the ethnic groups within space. For example, there is no distinction between a completely heterogeneous society in which two ethnic groups are split 55%/45% throughout the territory and another society in which those same ethnic groups are 100% dominant on different sides of the territory. Some authors, such as Fearon and Laitin, have corrected for the absence of religious considerations by creating religious fractionalization indices (Fearon and Laitin 2003b).

In the disaggregate approach, polygons can be made to identify the area that an ethnic group occupies to determine how many ethnic groups share a given state. This approach, while it can leave out some groups where specific areas are hard to identify, has the benefit of highlighting potential flashpoints within a country and more accurately comparing local data to ethnic composition. For my thesis, I chose to use this approach because it provided useful information on the locations of borders between ethnic groupings, which are often cited as a risk factor for conflict onset.

There are two primary sources of data on ethnic groups, their locations, and orientation: the *Ethnologue* dataset and Geo-Referencing Ethnic Groups (GREG) project. The *Ethnologue* dataset is the most commonly used dataset in existing disaggregated civil war studies and consists of maps of the locations of different linguistic groups throughout the world (Grimes 2000). The *Ethnologue* dataset is constructed starting in 1951 with "10 mimeographed pages and included information on 46 languages or groups of languages" (Lewis 2009). Over the past several decades, and with the input of a wide range of scholars, this compilation has grown into a complete printed and online compendium of language maps and data. Although several studies used these language maps as proxies for the locations of ethnic groups, most authors have noted that there is more complexity to the relevance than merely its language; an ethnic group's religion and other cultural traits can also be significant (Buhaug and Gleditsch 2005; Hegre et al. 2007; Rød and Buhaug 2007).

The Geo-Referencing Ethnic Groups (GREG) project, completed recently in 2009, has produced useable GIS maps of ethnic groups (Cederman et al. 2006, 2007; Weidmann 2008b). The Geo-Referencing Ethnic Groups project relies largely on data from the *Atlas Naradov Mira*, constructed by Soviet ethnographers in the 1960s (Bruk 1964), which produced comprehensive maps of the locations of ethnic groups around the world. The data from the *Atlas Naradov Mira* has been used for the construction of ethno-linguistic fractionalization (ELF) indices in country-year studies – Fearon and Laitin use it as a source of data for constructing their own ethnic fractionalization measure, and Collier and Hoeffler also use *Atlas Naradov Mira* as part of their polarization metric (Fearon and Laitin 2003b; Collier and Hoeffler 2004). Using *Atlas Nardov Mira* in a geo-referenced study – as it was designed to be used – should improve its analytic power. The latest version of the GREG data is a set of digital maps comprising 8969 polygons defining ethnic groupings worldwide. Each polygon can contain up to three ethnic groups. Worldwide, a large proportion of the polygons (7383 out of the 8969 polygons) contain only one ethnic group; there are 1552 polygons containing two groups; only 34 polygons contain three groups. (In Africa there are very few areas with three designated groups, so the third group has been ignored for the purpose of this analysis.)

Like aggregate ethnicity measures, the GREG data also has several notable shortfalls. First, it is somewhat out of date; as noted earlier, the source data was taken from Soviet studies in the 1960s. It would be facile to assume that ethnic distribution remains stable over time; some conflicts significantly alter the ethnic composition of states through ethnic cleansing or a mass migration away from conflict. . Indeed, as Weidmann demonstrates using a disaggregated analysis, conflict itself can have a meaningful impact on the ethnic composition of states, making it more difficult to accept the idea that ethnic balances remain unchanged in conflict (Weidmann 2008a). Given infinite resources, it would be ideal to compare these cases of ethnic fractionalization with revised data for each five-year period. Unfortunately, that is not achievable for my study – but any fluctuation in measurement of ethnic fractionalization over time would represent an improvement over existing studies. In the major aggregate and disaggregate studies, ethnic fractionalization is held constant over the entire span of the study from prior to 1970 until 1999 or later (Fearon and Laitin 2003b; Collier and Hoeffler 2004; Buhaug and Rød 2006). So by incorporating data from the 1961 and 1985 snapshots conducted by Roeder, it should be possible to at least provide some updated ethnic fractionalization information, even if only at the aggregate level.

A second shortfall of GREG and *Atlas Naradov Mira* is that it excludes ethnic groups that lack a defined geographic area. As a result, maps of countries with significant dispersed immigrant populations throughout the state do reflect the ethnic diversity of those countries. For example, the ethno-linguistic fractionalization index of the United States is 0.575 when calculated based on demographic data (Roeder 2001), but when a comparison index is calculated based on the geographic polygons identified for ethnic groups, it is less than 0.1 (Weidmann 2008b). The absence of dispersed groups from the GREG dataset is concerning, but it is be possible to include some aggregate measures, such as Roeder's ethnolinguistic fractionalization scores, to account for these differences.

Despite absence of dispersed ethnic groups, the concentrated ethnic groups are a reasonable focus for the data because they can have a much higher chance of being linked to conflict. At a theoretical level, this link between concentrated ethnic groups and violence seems reasonable because groups of people acting together engage in violence, while dispersed groups face problems of collective action. This theory is supported by the data from the Minorities at Risk (MAR) project. Toft's study using data from MAR concluded that concentrated ethnic groups were almost a necessary condition to for the onset of ethnic violence and that dispersed ethnic groups were almost incapable of engaging in collective violence (Toft 2003). Thus, the missing data on dispersed ethnic groups appears unlikely to explain instances of ethnic violence, although it may explain the absence of violence in some places.

A similar and related shortfall is the absence of data in GREG about the ethnic proportions within the mutli-group areas. For example, when Afrikaners and Englishmen are marked in sections of South Africa, there is no indication on whether the division between the two groups is 90%/10% or 55%/45% in the polygon defined by GREG. The absence of this information is regrettable, but it cannot easily be collected at this stage so this study works around the absent information.

GREG also neglects to identify some pertinent and well-known cleavages within larger ethnic groups. For example, almost all of Somalia is designated as the same ethnic group – Somali – even though inter-clan warfare has been cited as the source of considerable bloodshed since the fall of the government in 1991. Though central Africa, there are a large variety of smaller ethnic groups designated, and it is beyond the scope of this study to attempt to make individual determinations about whether those smaller groups should have been consolidated. Indeed, it may be impossible or even inappropriate to attempt to determine which ethnic cleavages are salient because it would rely on subjective judgment in many instances. It would be counter to the overall methodology of this study to make those determinations on a case-by-case basis, often relying on after-the-fact understandings of the conflict to make such determinations. A related shortfall of GREG is its division of large ethnic groups based primarily on national borders. This situation is most apparent in the case of African Arabs, who are divided into Egyptian Arabs, Sudanese Arabs, Libyan Arabs, Tunisian Arabs, Algerian Arabs, etc, by the *Atlas Naradov Mira* and thus by GREG. While there may be valid analytical reasons for dividing Arabs into subgroups, a quick review of the groups and the categories reveals that boundaries are national borders, rather than cultural differences. Dividing ethnic groups by political boundaries is particular challenge in the disaggregate context as it creates difficulty in assessing whether common ethnic groups exist on both sides of a political border, providing opportunities for safe havens and support. In aggregate studies, which assess every country independent of its neighbors' ethnic groups, this issue does not arise. Despite the many challenges of working with data from an outdated Soviet map, even if updated and digitized recently, the information from GREG is the most comprehensive geo-referenced information available, and the underlying data from *Atlas Naradov Mira* has been the primary source of information for ethnic measurements in other major comparative studies of civil war (Fearon and Laitin 2003b; Collier and Hoeffler 2004).

Since previous studies with GREG have largely used dyads rather than the full map of ethnic groups (Buhaug et al. 2006; Cederman et al. 2009), these studies do not clearly indicate the variables and data specifications to use when working with a full ethnic map, as

I do in my study. Also, the major prior aggregate studies do not provide clear guidance on the issue because aggregate and disaggregate studies can be particularly distinct in their approach to ethnicity and because different aggregate studies have focused on different data specifications. Based on the data available from GREG, this study will focus on two key variables to evaluate the relevance of ethnicity to conflict: 1) the distance of a grid cell to an ethnic group border, and 2) the density of ethnic borders in an area. For the purpose of this analysis, an "ethnic group configuration" is any distinct identification of ethnic groups in an area. So, for instance, in Sudan, one ethnic group configuration is Sudanese Arabs only while another would be Sudanese Arabs and Nubians. The border that divides these two ethnic group configurations is described as an "ethnic border" for this study. The original GREG maps measured ethnic groups as ending at either coastlines or in the uninhabited parts of the Sahara, but those borders have been removed for this analysis. Had the borders remained, then all groups living on the coast would have been next to an ethnic border, even if one was miles away. This situation is particularly clear in cases of the coast of northern Africa where Arabs had an 'ethnic border' with the Mediterranean but dominated the region and South Africa where the southern coast was marked as an 'ethnic border' even though none existed for hundreds of kilometers. (The exception to this general elimination of coastal borders is Madagascar. Because essentially the whole island is coded as homogenously Malagasy, eliminating the coastal border for Madagascar would have eliminated the island entirely from the ethnic map. While leaving the coastal border creates some counterintuitive data by making the ocean appear as another ethnic group, it would have also been counterintuitive for the western side of the island to be dramatically closer to an ethnic border – on continental Africa – than the eastern side of the island.)

Because all data must be rasterized to a single grid, assessing the map on the basis of ethnic group configurations rather than overlapping ethnic groups, allows for easier use of data. While there may be some clever methods for determining analytic outcomes from overlapping areas, particularly where one ethnic group is completed surrounded by another one, it is beyond the reach of the basic methodological design of this study to pursue those results. The information on the number of ethnic groups present was not incorporated directly because the range would have been very limited (0-3). Moreover, there are so few areas with the maximum number – three – of designated ethnic groups in Africa that including them would distort the data.

The distance to an ethnic border is the distance from a given grid cell to the nearest ethnic border, measured in the same method that distance to an international border was determined. The density of ethnic borders in an area was determined to provide some measure of the complexity of groups in an area since the number of groups in any configurations could only range to 3 at most. As a result, a kernel density function was used, in a very similar manner to the petroleum density variables, to generate a measure for the amount of ethnic borders in an area. The size of an ethnic group configuration overall provides a sense of its size regardless of political boundaries, and it is measured based on the number of grid cells that the configuration occupies. This variable does struggle because the level of granularity of the data in GREG is much higher than the 0.4 x 0.4-degree grid cells. As a result, some ethnic configurations are not considered because the territory that they occupy is so small. The size of an ethnic configuration within a country (in a given five-year period) is a ratio of how many grid cells are occupied by the configuration compared to the number of grid cells in the country. For example, the Xosa alone occupy 89 grid cells of 1159 cells in South Africa 1970-1974, so those cells are designated with 7.7% for that ethnic configuration. However, the Xosa and Englishmen occupy 5 of the 1159 cells of South Africa 1970-1974, so those 5 cells are marked with 0.4% for that ethnic configuration.

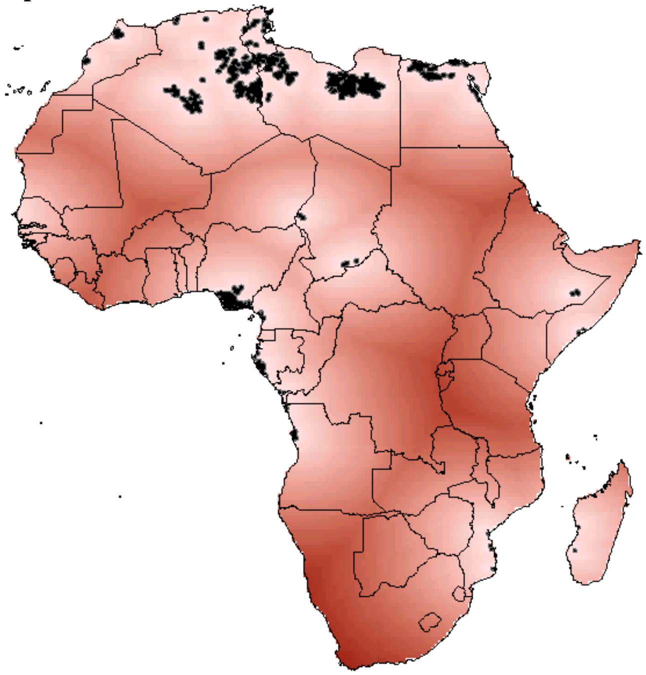
There are other variables that could have been considered – such as the roughness of the terrain occupied by the ethnic configuration or the distance of the centroid of the ethnic configuration to the state capital, but those variables are better assessed through a dyad study. The most recent publication on GREG is just such a study, and so no attempts have been made to duplicate GREG's work. Another distinction that could be made is to code for either the ethnic group that includes (or not) the state capital in its territory. Similarly, one could add information about which ethnic group is in power. Unfortunately, that information changes significantly over time, and, with a static ethnic map, it becomes impractical to attempt to measure its impact on conflict.

Natural Resources

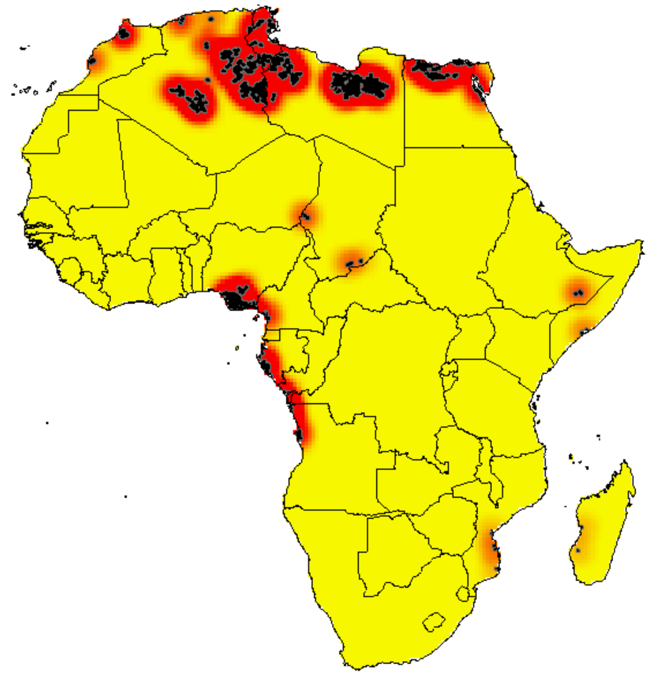
Identifying the relationship between natural resources and conflict has proven difficult within the existing country-year approaches. As is described in the literature review, aggregate studies have tended to examine natural resource revenues and exports, while disaggregate studies have tended to examine natural resource production and location In an effort to combine the interests of both aggregate and disaggregate studies in this project, I elected to include both measures. The revenues measure, specifically the portion of a country's total exports that come from primary commodities, is included in the economic information from the Collier and Hoeffler dataset (Collier and Hoeffler 2004). While there are difficulties with this variable, such as its inclusion of agricultural products and oil and the exclusion of non-petroleum minerals, it is a relevant variable and one that has been used in a range of studies in the field.

My study focuses on contributing more to the understanding of how the location of natural resource deposits impacts conflict. To test these hypotheses, PRIO has developed two geo-referenced datasets, one for the distribution of petroleum resources and the other for the distribution of diamond resources (Gilmore et al. 2005a; Lujala et al. 2007). The petroleum dataset identified 98 onshore petroleum deposits in Africa. Unlike the diamond database, only 29 (29.6%) of the deposits lack a date of discovery, and the deposits that lack a date of discovery appear in groupings of several other dated deposits. Thus, it is possible to construct a straight line distance from each grid cell to an onshore petroleum deposit for each of the five-year periods without losing a significant amount of data by excluding those sites with no production date. Also, it is more reasonable to assign the dates of discovery to petroleum and not to diamonds because wealth from petroleum cannot be accessed without considerable infrastructure, meaning that any locally known deposits would not have held significant value until they were exploited starting with their date of discovery. On the other hand, diamond deposits could be locally known and exploited by artisanal miners without any major discovery or the deposit becoming too widely known. (As with the diamond data, this straight line distance for the petroleum dataset is constructed without consideration of national borders.)

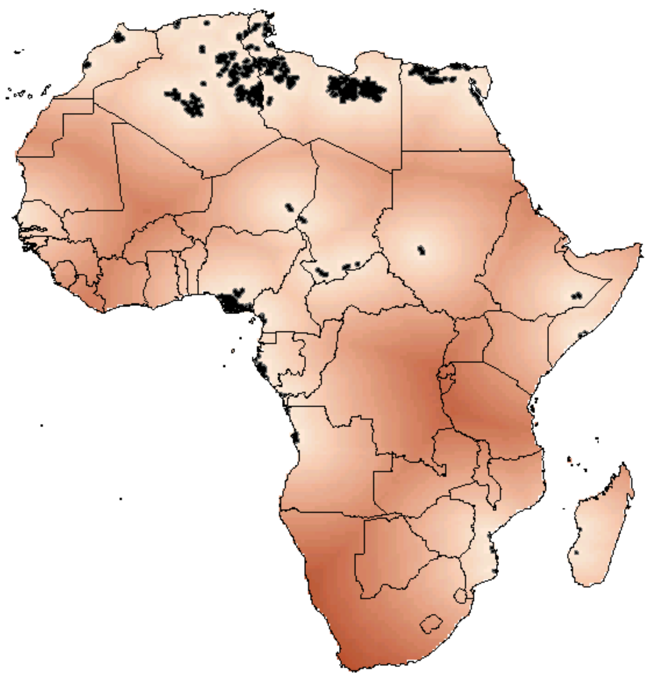
Figure 8. The tables below show the distance to and density of onshore petroleum in the six periods covered. Darker colors indicate higher values, and the black sections mark the shapes of the petroleum deposits.

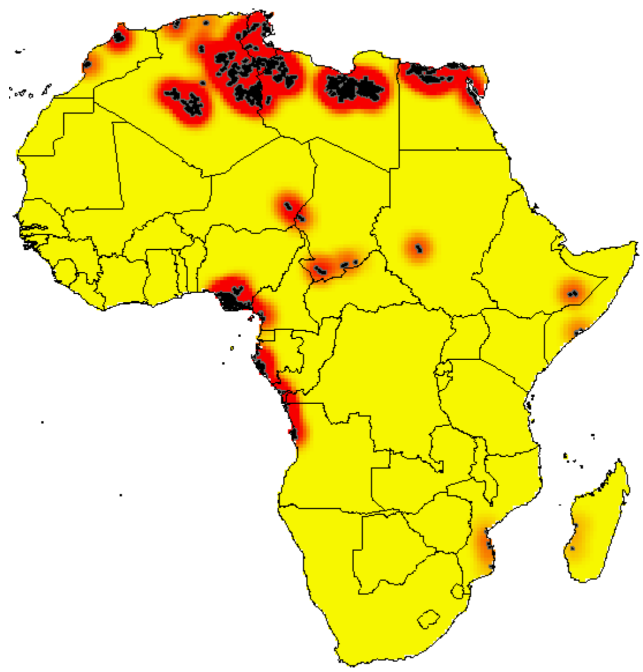


1970-1974 Distance to Onshore Petroleum



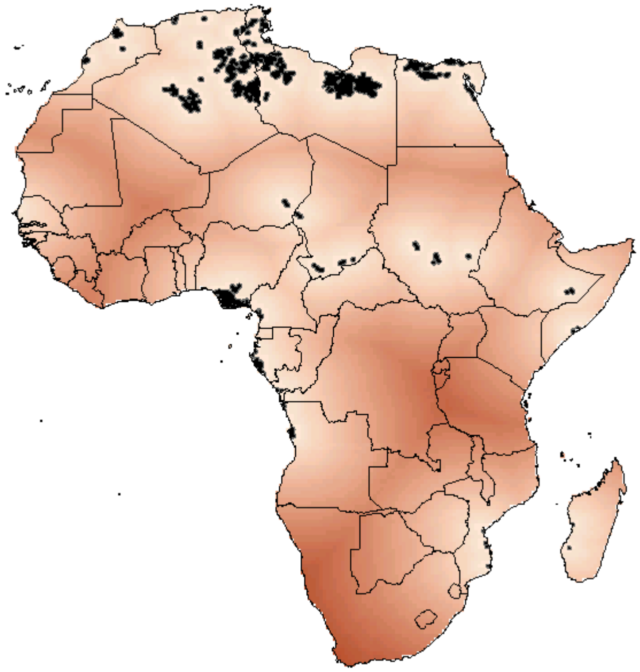
1970-1974 Density of Onshore Petroleum

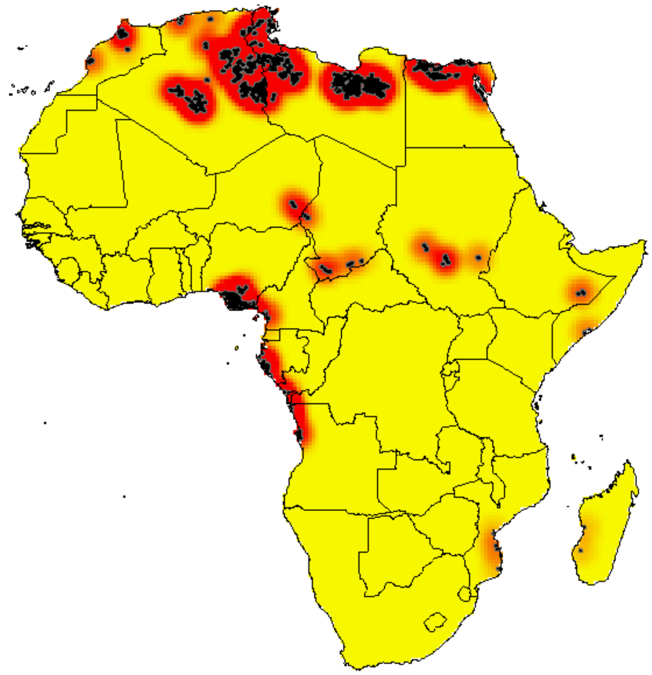




1975-1979 Distance to Onshore Petroleum

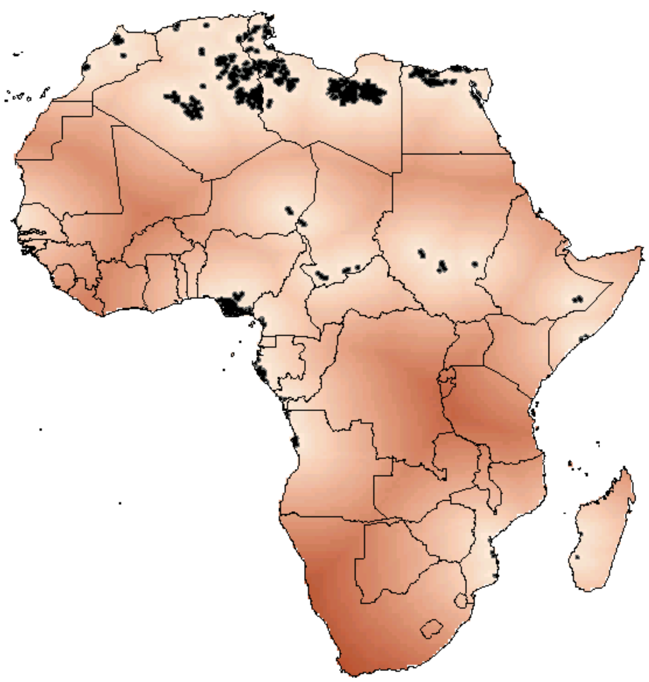
1975-1979 Density of Onshore Petroleum

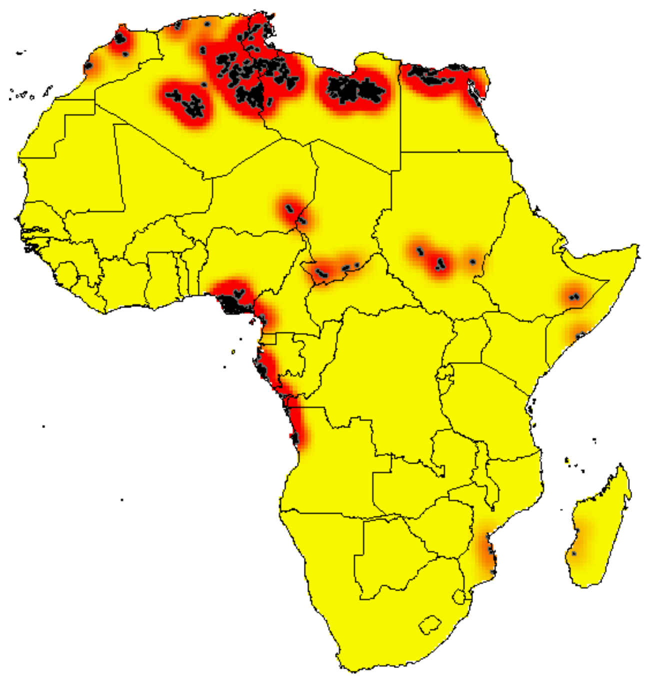




1980-1984 Distance to Onshore Petroleum

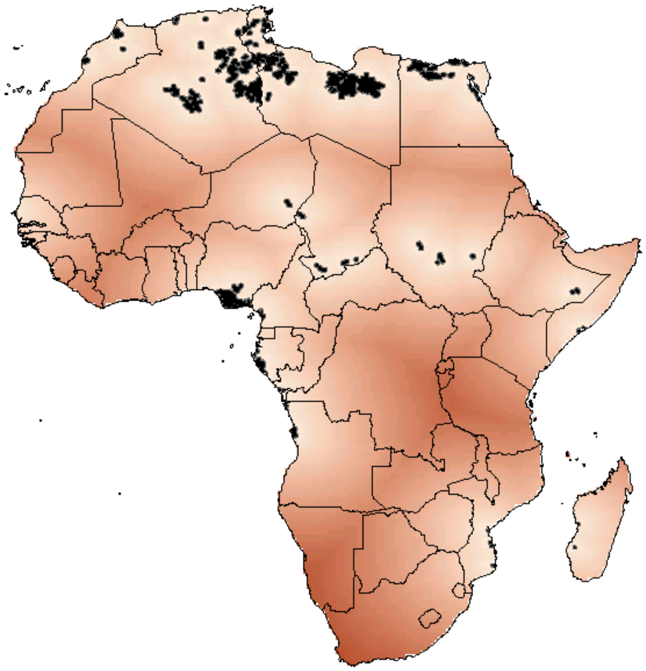
1980-1984 Density of Onshore Petroleum



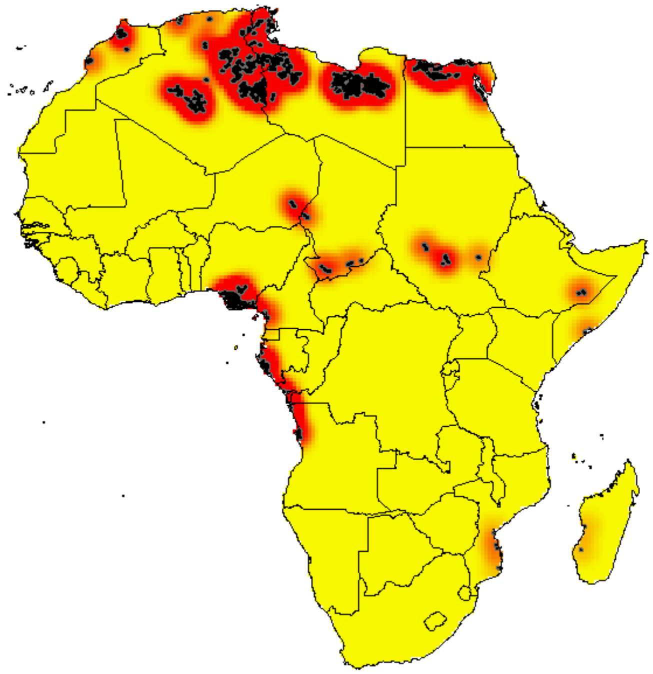


1985-1989 Distance to Onshore Petroleum

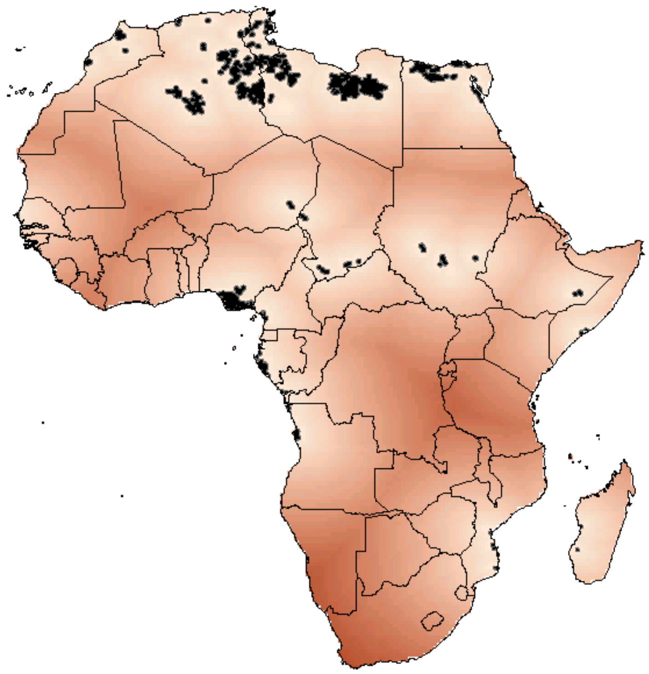
1985-1989 Density of Onshore Petroleum



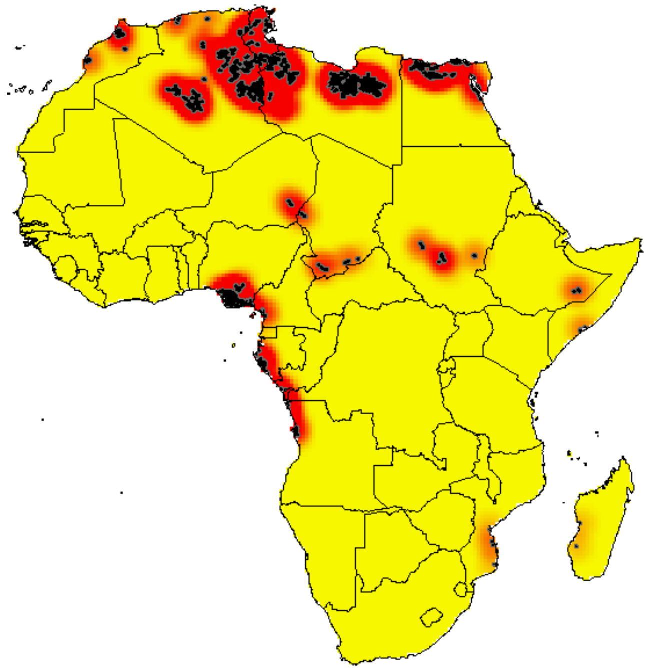
1990-1994 Distance to Onshore Petroleum



1990-1994 Density of Onshore Petroleum



1995-1999 Distance to Onshore Petroleum



1995-1999 Density of Onshore Petroleum

The PRIO diamond dataset includes information on 1176 different diamond deposits worldwide. Only 81 locations (comprising 496 deposits) are marked as lootable diamonds with known production and only 25 locations (comprising 90 deposits) are non-lootable diamonds with known production. The breadth of the global diamond dataset naturally means that it does not have depth of information on some sources. Thus, all diamond deposits designated as secondary diamonds were used – diamonds that could potentially be looted – even though some may be more easily lootable than others based on their location and level of development. This set of secondary diamonds equated to 306 deposits in Africa. Unfortunately, the dataset of secondary diamonds does not have enough information to restrict data on secondary diamonds into five-year periods without excluding 206 (67.3%) of the African deposits for which there is no known date of discovery, so all secondary diamonds were included in all time periods, regardless of their date of discovery. (Including secondary diamonds without known dates of discovery should not skew the analysis, as secondary diamonds are more accessible. Local residents might exploit the deposits even if there was not yet a formally recorded "date of discovery" with Western sources that catalogue mineral deposits.)

I generated two variables to evaluate the effect of diamonds on conflict. The first was the straight line distance from any grid cell to the nearest secondary diamond deposit. As measured for this study, the straight line distance for diamonds did not take into account the borders of countries. Buhaug and Rød take these borders into account, arguing that insurgent groups in neighboring countries do not generally seek secession because diamonds on the other side of the border. While it is true that the existence diamonds in a neighboring state may not be the grievance of an insurgent group, case study evidence suggests that they may serve as a funding source (Smillie et al. 2000; Malaquias 2001; Ross 2002, 2006). Also, when measuring diamond distance only within the country, the problem arises of determining the diamond distance for states with no diamonds. Buhaug and Rød assign an arbitrary high value to the grid cells in these states, but it seems more reasonable (and more accurate) to simply ignore the borders when assessing distance to diamonds, particularly as such borders are often porous. The second variable generated to measure the impact of diamonds is the density of diamond deposits in a given area. When evaluating the map of secondary diamond deposits generated by plotting the locations of the deposits, it was clear that there were some areas that were markedly denser than others, but that this distinction was not being captured by the simple straight-line distance variable. Using the geo-referenced diamond dataset, a density measurement was taken for the number of diamond deposits in a set radius from any given grid cell. While it would have been preferable to include a variable that measured the density of diamond deposits by their monetary value, rather than simply the number of deposits, such information was not available. (More detail on the methods used for density calculations is available in the spatial lag section of this chapter.)

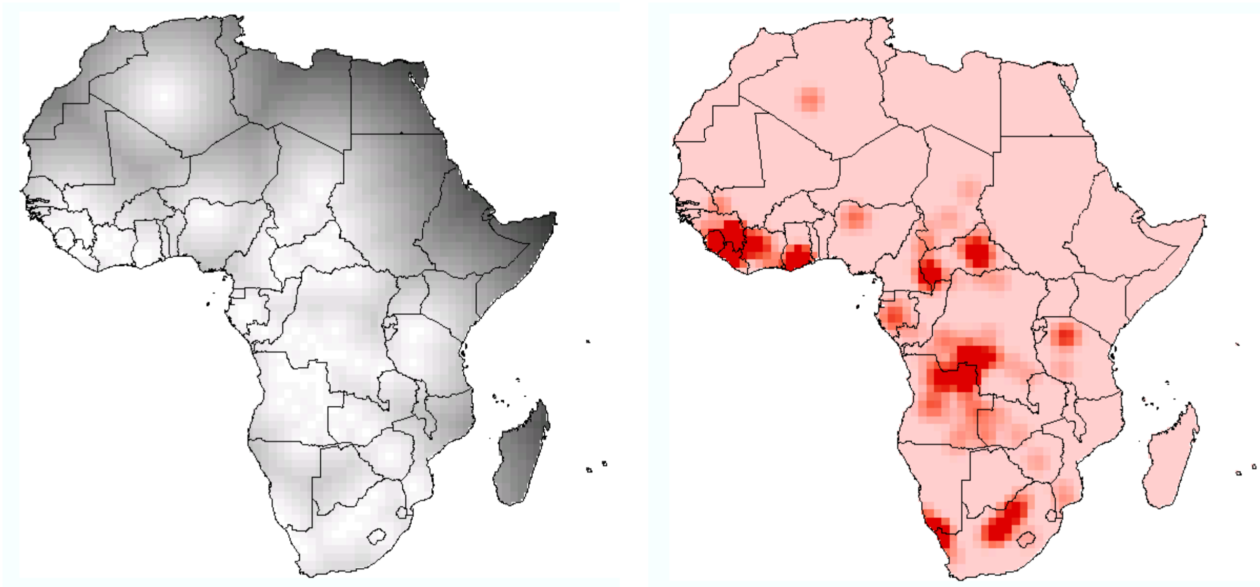


Figure 9. The map above on the left shows the straight line distance variable from secondary diamond deposits and the map on the right shows the density of secondary diamond deposits. Darker colors indicate higher values in both cases. The data on secondary diamond deposit locations was drawn from DIADATA, and the distance and density mapping is my own.

To date there are no comprehensive datasets linking other resources, such as timber, narcotics, and certain agricultural resources (e.g., cocoa) for instance to conflict. While the absence of these datasets is a concern for getting the full set of natural resources that can be linked to armed conflict, alluvial and kimberlite diamonds are by far the most profitable and, in the case of alluvial diamonds, also the most easily extracted non-fuel mineral resource. Thus, while we would expect there to be an association between conflicts and financing from a wide range of natural resources, diamonds appear to be the most significant source of financing (Le Billon 2005). The importance of diamonds is not surprising – in terms of "lootability," diamonds are by far the most efficient common mineral. In 2000, the average price of diamonds was approximately $355,000 per kilogram (or $71 per carat), while gold was valued at approximately $9000 per kilogram ($280 per troy ounce), less than gold by a factor of forty (Amey 2000; Olsson 2005). Despite this massive difference in "lootability," other minerals are sometimes cited as the sole natural resource driving a conflict (although there may also be significant political and economic factors). For instance, Ross cites the 1960-1965 rebellion in Katanga in southern Democratic Republic of the Congo (then Zaire) as driven by copper (Ross 2004b), as well as the political instability that occurred when Zaire gained its independence from Belgium in 1960. At the time that the rebellion started – 1960 – copper prices reached a high of about $0.33 per pound, but gold was valued at $35.00 per troy ounce – roughly 1546 times more valuable per pound than copper (Callaway et al. 1961; Ryan and McBreen 1961). So, although diamonds and oil have been some of the most often identified natural resources associated with conflict (Ross 2004a), making them necessary to consider in order for this study to be comparable to other work in the field, there is evidence that minerals other than diamonds can be relevant even if they are markedly less valuable by weight.

While data on natural resources such as crops, timber and narcotics are not available, there is a geo-referenced database of mineral resources – the United States Geological Survey Mineral Resources Data System – which includes a wide range of minerals, including gold, silver, copper, coltan (columbium), uranium, and chromium, among others – and this dataset has not been used in previous studies of civil war (*Mineral Resources Data System* 2005).[12] The dataset includes just over 2700 mineral deposits in Africa (out of over 305,000 worldwide), ranging from “occurrences,” about which little information is available, through to entries with complete information about dates of discovery, production, and the types of minerals being developed. Initially, this data source appears difficult to use because only a portion of the relevant entries have dates. For instance, of the 452 entries for gold deposits in Africa, only 96 have dates for their first

[12] Searches of published articles prior to 2009 found no civil war articles that used this dataset for aggregate or disaggregate analysis of the role of minerals in conflict.

production. Unlike the DIADATA dataset, which is a focused dataset designed for use on civil war studies (Gilmore et al. 2005a; Lujala et al. 2005), the USGS dataset requires researchers to make some judgments about which minerals to include and how to include them.[13] Based on the decisions made in the construction of DIADATA, "occurrences" were held to be enough information about a mineral resource to include it. In the codebook for DIADATA, an "occurrence" is defined as "any site with known activity, meaning production (either commercial or artisan) or confirmed discovery" (Gilmore et al. 2005b). The USGS dataset uses the same terminology to define the least significant state of development for a mineral resource, but this status often means that nothing is known about it. If the standards for DIADATA's use in its first publication are maintained for USGS, the number of entries in USGS is reduced to 2191 entries by taking only data about some information is known beyond the location and type of mineral (a discovery or production date, an amount of the deposit, or an operation type) and eliminating any deposit known to be too

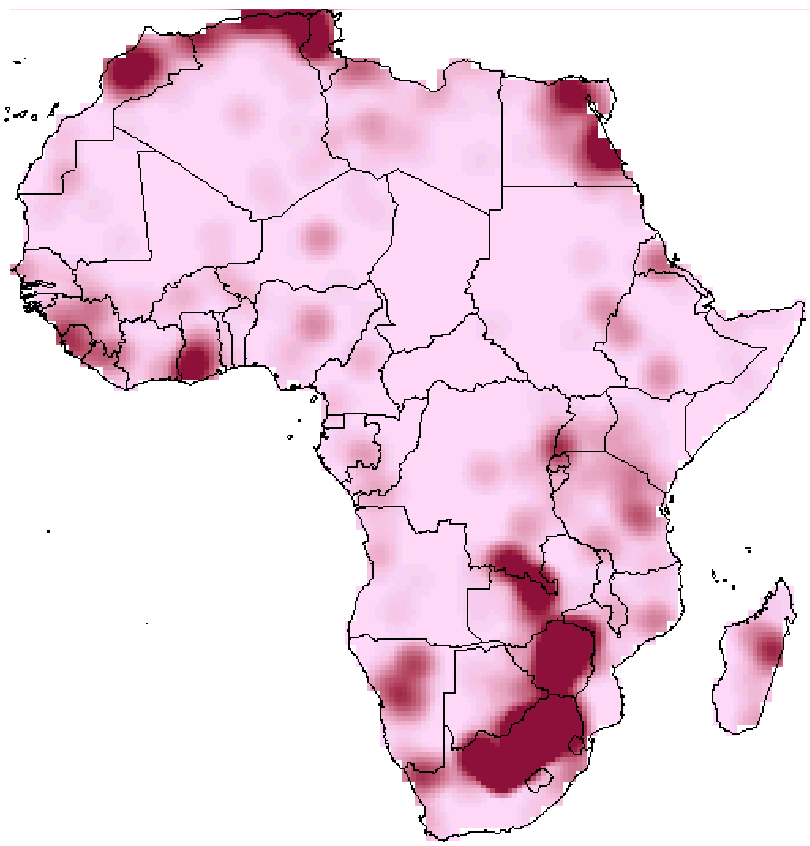


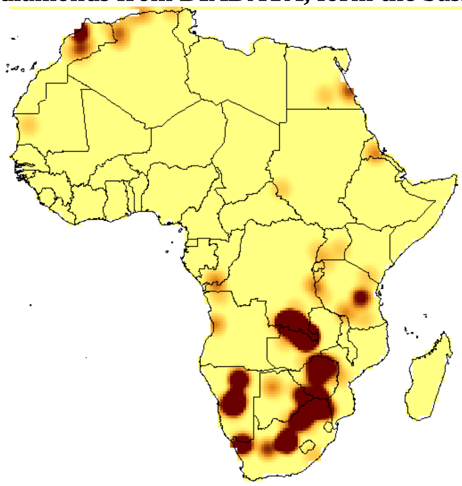
Figure 10. The density map above shows the composite of the mineral data used from the US Geological Survey and diamond data.

[13] The DIADATA dataset does not include the US Geological Survey in its source listings; it is clear that DIADATA is a much more complete set of information on diamonds. The USGS dataset has 29 diamond deposits recorded – DIADATA includes 52 in Angola alone.

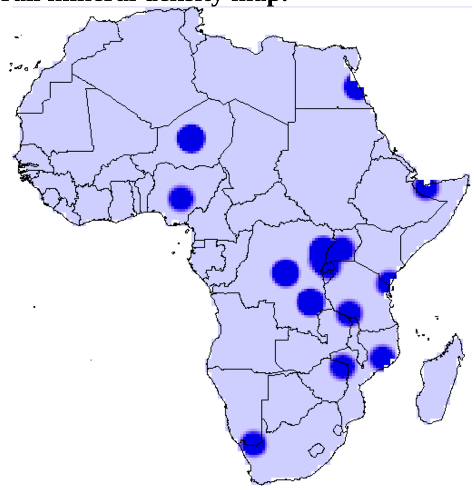
small for production. Then, taking only the deposits for which a mineral of interest is listed as a primary mineral (rather than a secondary or trace mineral present in a larger deposit), the shapefiles for the dataset used in this project are constructed. This process leaves 407 deposits for gold, 217 deposits of copper, 151 deposits of chromium, 109 deposits of iron, 79 deposits of manganese, and 71 deposits of platinum (almost all of which is South Africa) – beyond platinum, almost every other mineral has so few deposits that it is difficult to suggest that it serves as a general variable and not just a dummy variable for one specific region.

There are very few other minerals in the USGS dataset that have been suggested as associated with multiple conflicts. The one notable exception to this trend is coltan (which is measured by including any deposit of niobium, columbium, or tantalum), which only has 18 deposits that meet the standards, but it appears in 10 different countries and has been claimed to fuel conflict in the Congo ("Report of the Panel of Experts on the Illegal Exploitation of Natural Resources and Other Forms of Wealth of the Democratic Republic of the Congo" 2001; Montague 2002; Jackson 2003). One reason why coltan is an attractive resource to rebels is that it can be processed to a great extent by local groups without sophisticated equipment, although the importance of coltan may fall outside of the scope of this study since its major price spike, reaching as much as $380 per pound, occurred in December 2000, due to a massive increase in coltan demand from the release of coltan-heavy PlayStation 2 (Vesperini 2001).

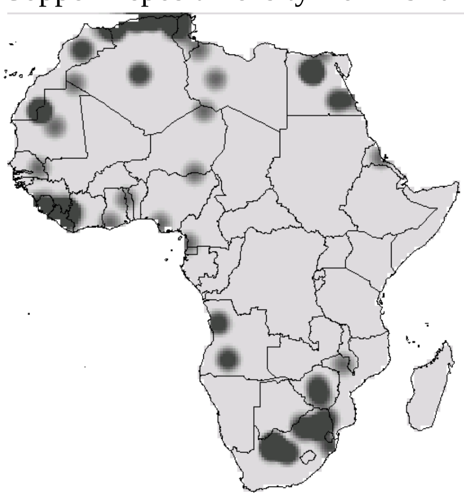
Figure 11. Below are the density maps for copper, coltan, iron, gold, manganese, and chromium that were developed from data from the US Geological Survey. These density maps, along with data on diamonds from DIADATA, form the basis of the overall mineral density map.



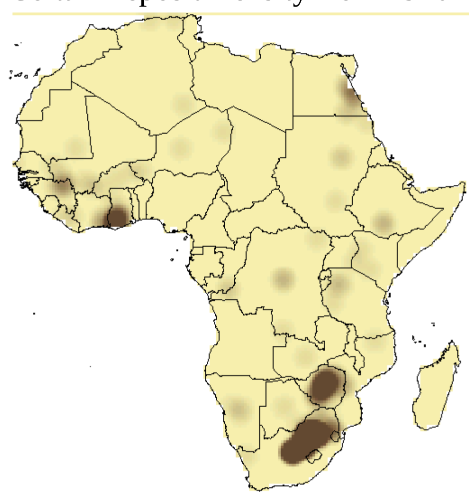
Copper Deposit Density from 1970-1999



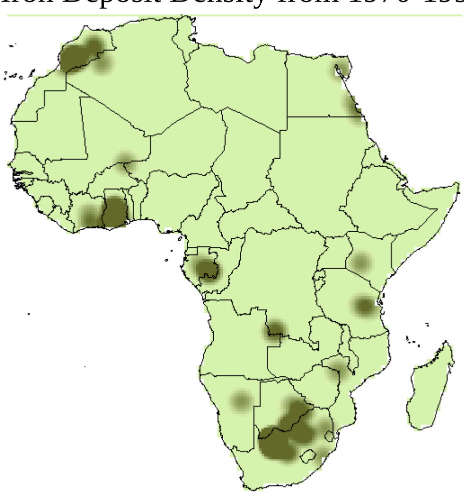
Coltan Deposit Density from 1970-1999



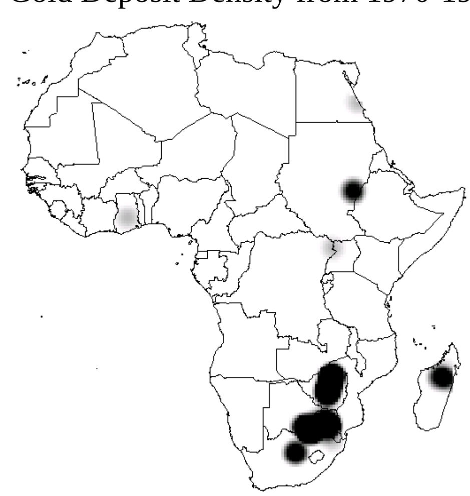
Iron Deposit Density from 1970-1999



Gold Deposit Density from 1970-1999



Manganese Density from 1970-1999



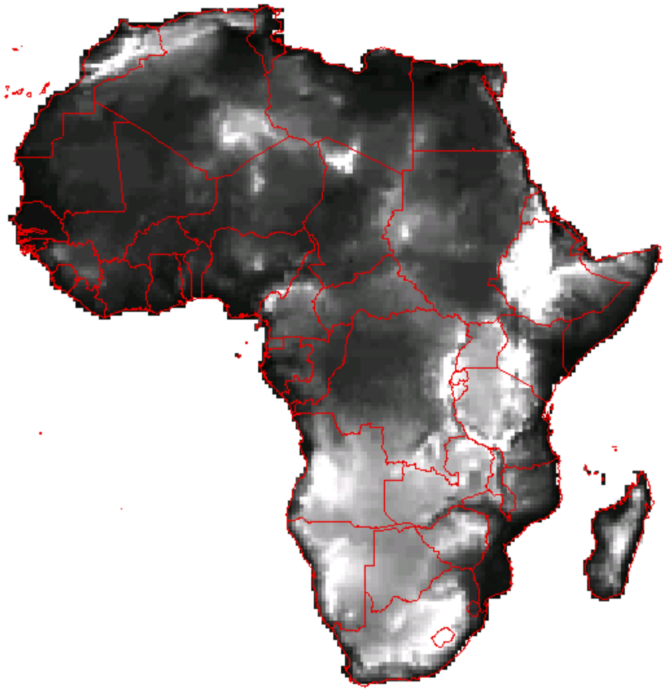
Chromium Deposit Density from 1970-1999

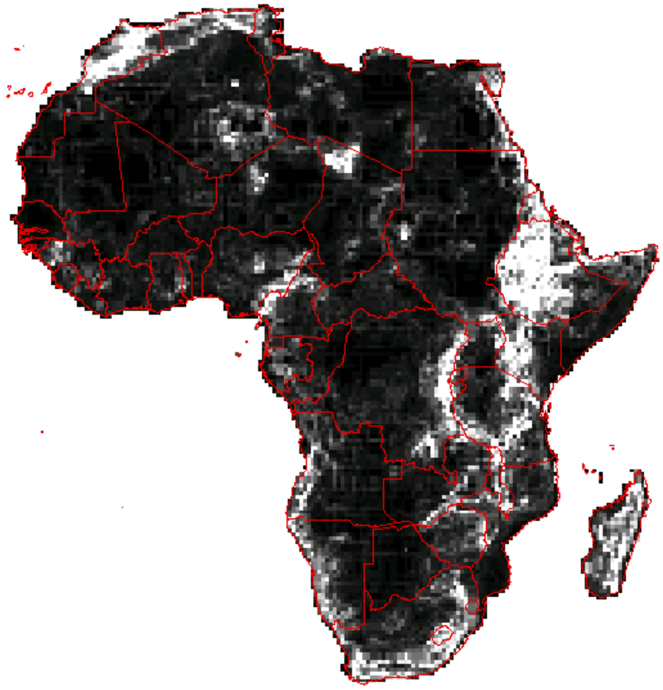
There was no readily available set of standards for comparing minerals – whether against their "lootability" or any other standard – so this study will simply use the same fairly crude tool of measuring the proximity and density of mineral deposits. While this data construction is far from a perfect approach, it will add some more information to the evolving debate over the role that non-fuel minerals play in conflict.

NOTE: This thesis does not include narcotics data even though it is a noted element in many civil war studies (Ross 2002; Collier et al. 2004; Fearon 2005). In most of these studies, drugs are either noted without being quantified or they are considered with dummy variables, since there is little credible data on the drugs situation in Africa. (For instance, only 24 of 53 African countries responded to a mandatory survey on the drugs and crime situation in their countries from the UN Office of Drugs and Crime in 2004 (Sinha 2009).)

Mountainous Terrain

In prior aggregate studies, roughness of terrain was constructed using expert judgments and interpolations, rather than any computer-aided assessment of the terrain. For example, roughness of the terrain from mountains was measured simply as a matter of the percentage of a country's territory that was covered by designated "mountainous terrain" by an expert, A.J. Gerrard (Fearon and Laitin 2003b; Collier and Hoeffler 2004). When Gerrard's index did not cover a country, Fearon and Laitin interpolated the level of mountainous terrain from the mid-point of the highest and lowest points in a country cited in the CIA World Factbook (Fearon and Laitin 2003b). Disaggregate studies have used similar approaches, focusing largely on a designation of "mountainous terrain" based on data from UNEP (Buhaug and Rød 2006). While a designation for the percentage of mountainous terrain is useful, a general percentage does not indicate the severity of the mountains, nor does it provide a clear gradation between completely flat ground and extremely mountainous terrain within a country. This disaggregation would allow for hills and major rifts to also be recognized as roadblocks (literally and figuratively) in the ability of governmental regimes to project power into territory. (The use of this slope and elevation data to measure the distance from state capitals to outlying areas is explained in greater detail in the section on capital distance measurements.) Given the difficulty that can come with surviving in the mountains on limited food and water supplies and in particularly hazardous conditions, it may be that mountains are not in fact particularly desirable safe havens for rebels, as has been argued by Fearon and Laitin (Fearon and Laitin 2003b). In fact, it may be that rough terrain that provides havens without becoming hazardously mountainous is the most conducive to insurgent safe havens and therefore conflict.





Mean Elevation in Africa Elevation Range (by Grid Cell) in Africa

Figure 12. The two images above show the average mean elevation and the average range in elevations by grid cell in Africa. This data was collected from a more refined image from GTOPO30 and then I processed the data using ArcGIS 9.2 for my thesis.

Thus, two alternate measures for roughness of terrain are explored in this project, drawing on data from the US Geological Survey's GTOPO30 projections, which is a global digital elevation model (DEM) spaced on a horizontal grid of 30 arc-seconds for each cell (about 1 kilometer wide) (U.S. Geological Survey 2009). The first measure of roughness of terrain is the percent of slope from elevation changes in each cell. Because it is the percentage, and not the degree, the variable has a potential range of 0 to infinity, as it is the tangent of the angle of slope. There were many flat areas around the Sahara and the highest slope was close to 18% (or just over 86 degrees) in an area in the Great Rift Valley in Kenya. The method for gathering the slope data was to perform the slope calculation on the smaller resolution elevation map and then aggregate up to an average for the 0.4 x 0.4-degree grid cells that are the resolution for the overall study. The second measure for the roughness of terrain was the range of elevations from the 30 arc-second square cells presented in a larger 0.4 x 0.4 degree grid cell. The minimum was 0 meters – a perfectly flat cell – and the highest range was 2911 meter, located in Tanzania, near the border with Kenya. This combined method of gathering the range and the slope is particularly helpful in the Rift Valley where mostly flat cells can have precipitous drops in them, generating wide ranges of elevations but fairly level slopes. The two measures have a correlation of about r = 0.6, indicating that both specifications were not heavily overlapping.

Forested Terrain

Begins on next page due to formatting issues.

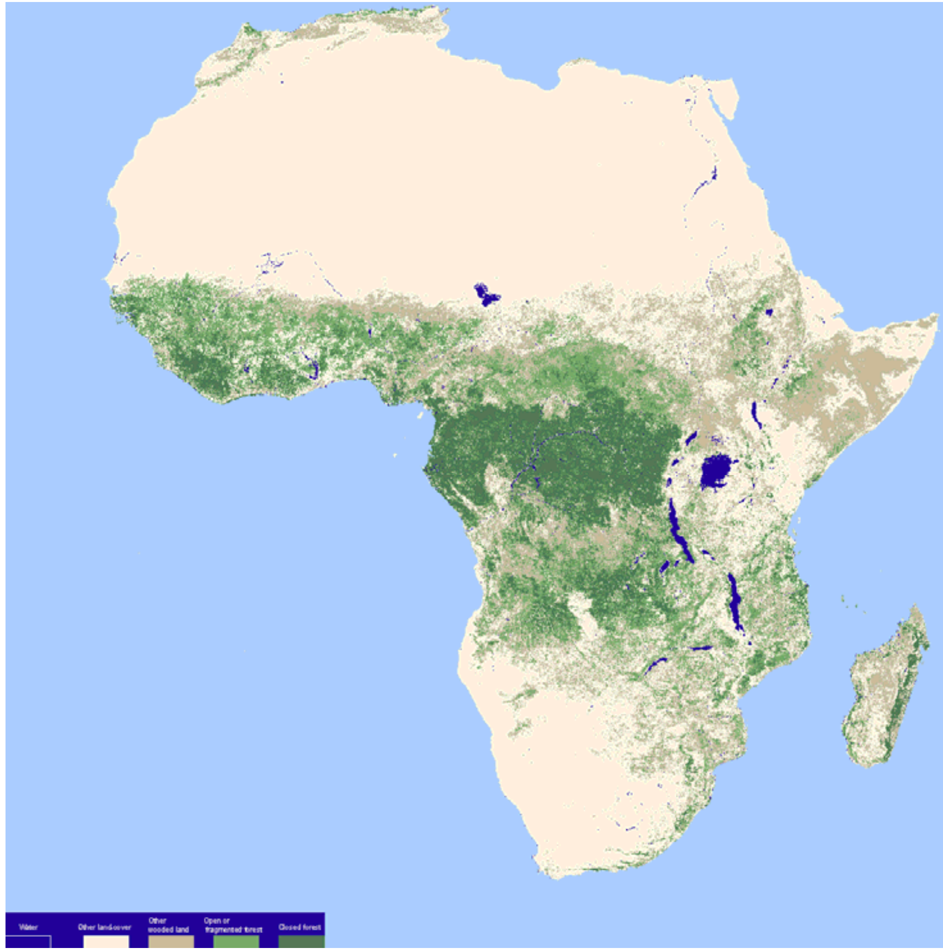


Figure 13. The map above is the processed satellite image of land cover provided by the US Geological Survey to UNEP-GRID. For my thesis, I reclassified the categories provided in the satellite image and then rasterized the image at a coarser resolution to produce the forest cover data.

Similar to Buhaug-Rød, data for the density of forested terrain was drawn from UNEP-GRID for forestry data from satellite images circa 1995 provided by the US Geological Survey. Since there were no comparable forest images for earlier periods, this forest data was used for all time periods. While this limitation is not ideal, particularly given some arguments that deforestation can lead to conflict (FAO 2005), the only other option was to ignore forest data, which was determined to be unacceptable given the numerous reports of rebels using jungles and dense forests as cover, particularly in Uganda and eastern Congo. The forest cover image from UNEP-GRID provided four categories for forest data - Closed Forest, Open or Fragmented Forest, Other Wooded Land, and Other Land Cover (*Global Forest Cover* 2000). On the basis of this data, a scale was established from 0-3 ranking the density of the forest. While this index does not directly correspond with the percentage of forest cover of any one type, it does provide a more accurate gauge of the "roughness" of the terrain from

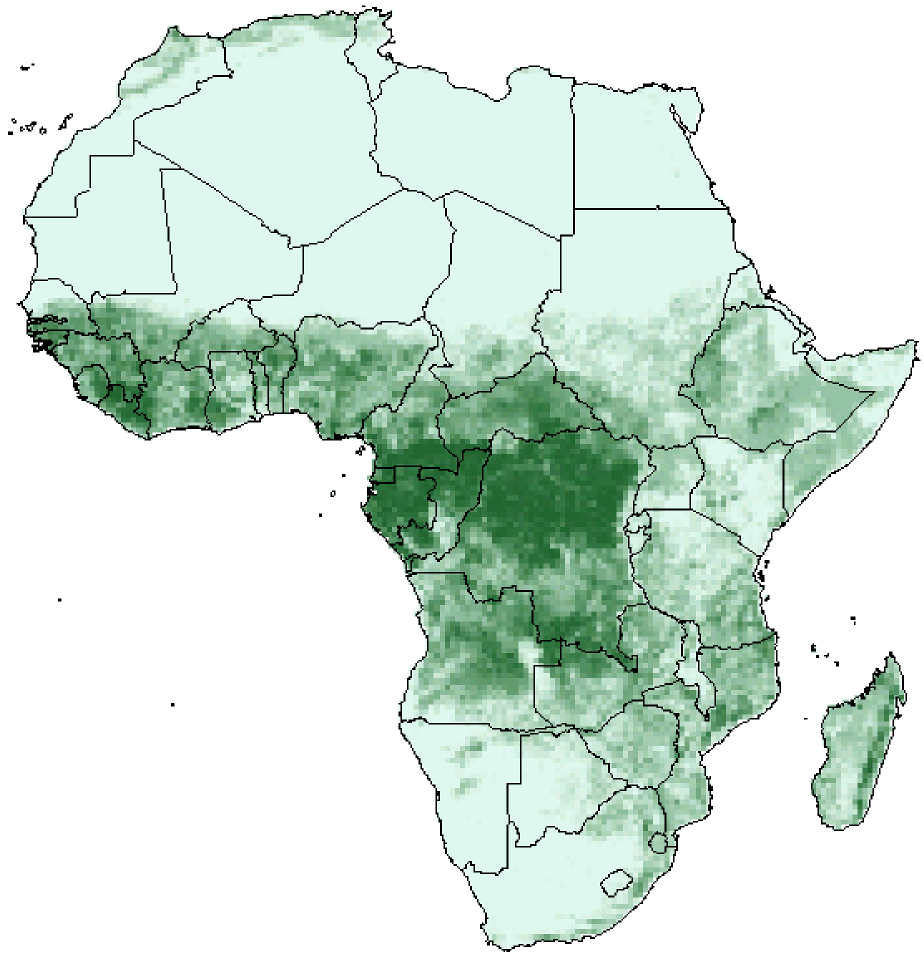


Figure 14. The map above shows the forest density that was used for all periods because information for each five-year period was not available. Darker values indicate higher values.

forest than only taking into account the densest forests as had been done in previous studies (Buhaug and Rød 2006; Collier and Hoeffler 2004; Fearon and Laitin 2003b). The values are skewed - the range of values runs fully from 0-3 with a mean of 0.657102, but a median value of 0.21875 and standard deviation of 0.845407. This skew appears to be due largely to the 6,638 grid cells in the Sahara that received a score of 0 for their level of "forestedness." For the purposes of statistical analysis, it would certainly be preferable to have a consistent, tangible measure that scales from completely barren terrain to the most densely forested terrain possible. As the variable is constructed for this project, it is little more than a general direction of "forestedness." Unfortunately, it will be difficult to interpret regression results in terms of adding "x trees per square kilometer" or "x percent more forested terrain in a grid cell." This disadvantage was considered, but given the lack of strong data connected with tangible units, the 0-3 gradient was used as a directional measure for the density of the forest.[14]

Conflict Data: Shifting to Polygons

Developing the data for conflict – the dependent variable in my study – came from a variety of sources. The key improvement in the tracking of conflict is shifting from an aggregate measure of conflict that assesses the entire territory of a state as experiencing a conflict regardless of the conflict's true size and location to conflict data in the form of conflict polygons that provide a more precise measure of exactly where a conflict occurred in a state in a given year. The data used for these conflict was developed by Jan Rød for his study with Halvard Buhaug on African civil wars (Buhaug and Rød 2006). As discussed in the literature review, conflict polygons are the latest step in a series of improvements – from conflict center points, to conflict circles, to ultimately conflict polygons – that refined the method for defining the shape of a conflict in space (Gleditsch et al. 2002b; Buhaug and Rød 2006). Each conflict polygon is also defined by its duration, marked by a start and end year, so it is possible to place the polygons in the relevant five-year periods for the data analysis.

The Buhaug-Rød conflict polygons notably does not take into account varying intensity of the conflict over the space of the polygon, changes in the center of the conflict over time, or the duration of the conflict overall (Hegre and Raleigh 2005). These issues are noted by Hegre and Raleigh as the basis for using the Armed Conflict Location and

[14] An interesting common metric for the thickness of a forest could be the amount of time required to walk 1 kilometer, but obviously gathering data on such a variable would be impossible for an entire continent.

Events Dataset (ACLED), which is a dataset developed to cover the conflict events during twelve civil conflicts in eight central and west African nations between 1960-2004 (Raleigh and Hegre 2005). In total, ACLED includes 4,145 "battle events," which includes battles, bases, rebel presence, and instances of territory transfer. The tracking of individual battles, while a tempting level of detail to reach, does raise questions about becoming *too* detailed. Because of the pinpoint specificity of the location of battles in this dataset, Hegre and Raleigh proceed to use grid cells of 8.6 x 8.6 km for their subsequent study (Hegre and Raleigh 2005). In studying a phenomenon that is as complicated and messy as conflict, the level of detail presented in the ACLED battle events increases the chance of presenting highly significant outcomes when a greater degree of uncertainty is probably warranted. Moreover, while the ACLED dataset is considerably more specific than the conflict polygons developed by Rød and Buhag, it covers a limited swath of territory in central and western Africa and expanding it would be extraordinarily time-consuming. Thus, while it is not ideal to accept uniform intensity across the entire conflict polygon, I decided that it was the only reasonable option at this stage.

One of the challenges of shifting to polygons for measuring both onset and occurrence is how to measure the spread of conflict. When conflicts are measured merely at the aggregate state level, the spread of the conflict within the country is not considered. With disaggregate conflict polygons, though, the spread of the conflict creates open questions for how to code the onset and occurrence of conflict correctly. In other words, if a conflict starts in the state capital and isolated urban environments, as occurred in Mozambique in the late 1970s, but then the conflict expands to cover the majority of the territory of the state, how should it be coded? Does each grid cell experiencing conflict for the first time get scored as having an onset even if it is not the onset of the core conflict? For my project, I opted to code only the true onset of the conflict in the first relevant year. In some cases, this meant that there is a slight mismatch between the onsets and occurrence in any given period (as in the case of Mozambique, for instance).

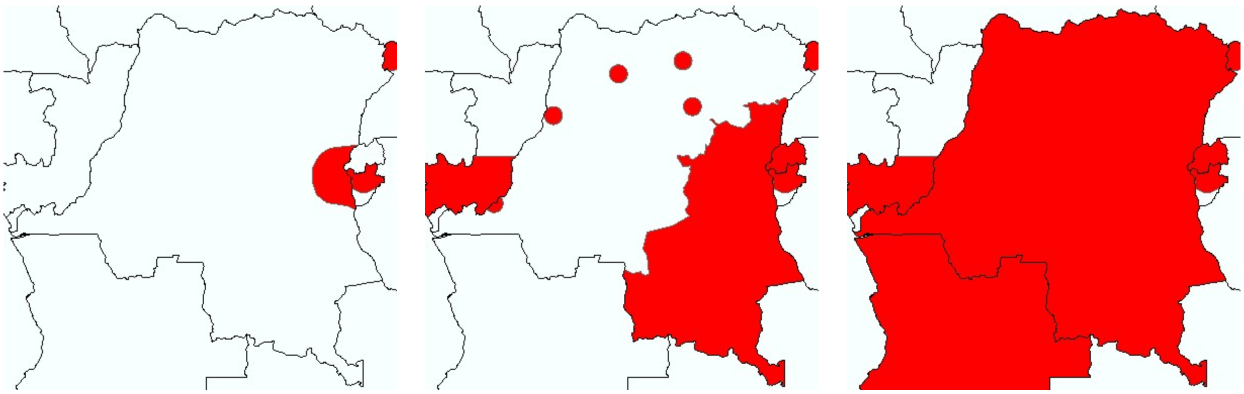


Figure 15. From left to right shows the expansion of the conflict in DRC as measured by the Buhaug-Rod conflict polygons in 1996, 1997, and 1998. When judging when and where onset occurred, the expansion of the conflict creates a challenge for coding disaggregate studies with multiple time periods. For this case, I opted to mark an onset in 1996 while marking the remaining territory as a conflict occurrence. While this creates some challenges for the dynamic logit analysis, it is a more accurate method of coding the geographic onset of the DRC conflict.

These mismatches create a problem for the dynamic logit approach to analyzing occurrence, but the measuring with the alternative approach– with rolling onsets for grid cells absorbed in the conflict – would have masked the true root of the conflict onset location. This approach is not perfect, but, based on the literature reviewed, there does not appear to be a clear correct way to code and model the many starts, restarts, and reconfigurations of civil wars.

Conflict Data: Distinguishing Territorial and Governmental Conflicts

One of the central areas of investigation for this thesis is whether the correlates of conflict are meaningfully different based on the aims of the conflict actors. In specific, the conflict data used distinguishes conflicts into two categories: one category where the non-state conflict actor aims to capture control of territory, whether for autonomy or outright secession, and a second category where the non-state conflict actor aims to capture the government of the state. The Uppsala/PRIO Armed Conflict Database (ACD) created the categories that are used in my study. The ACD differentiates conflicts into two primary forms of incompatibility:

a) "incompatibility concerning government," which the codebook defines as "incompatibility concerning type of political system, the replacement of the central government, or the change of its composition;" and,

b) "incompatibility concerning territory," which the codebook defines as "incompatibility concerning the status of a territory, e.g. the change of the state in control of a certain territory (interstate conflict), secession or autonomy (internal conflict)" (Gleditsch et al. 2002b; Gleditsch et al. 2002a; Harbom 2009).

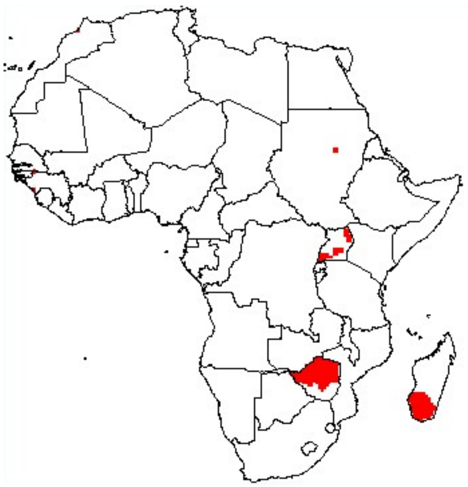
The ACD codebook states that the parties themselves – one of which must include the government – must articulate the type of incompatibility themselves to receive a classification as "territorial" or "governmental." The classification is open to criticism that it is either not sufficient to cover all the possible distinctions in aims, that coding may be based on false claims by the conflict actors, or the nature of the incompatibility may change over time. I reviewed these concerns, but I decided that they were not sufficient concerns to invalidate this approach for my study.

In regards to the first instance – that "territorial" and "governmental" distinctions are not sufficient – the original creators of the ACD note that problem but do not provide a ready solution for it other than excluding the cases. In a footnote of the article by Gleditsch et al., the original creators cite the example of how to classify the 1998 Al-Qaeda bombings on US embassies in Kenya and Tanzania, which occurred when the stated goal of Al-Qaeda was to change US policies in the Middle East, which is not covered by the definition of armed conflict" (Gleditsch et al. 2002b). According to the ACD authors, "although those attacks meet the criterion of more than 25 battle-related deaths, the 1998 events are not included in the dataset" (Gleditsch et al. 2002b). This example targeted foreign assets and really does not appear to be classified as part of any internal civil conflict, so it is not relevant to this project. As a review of similar incidents occurs, though, reveals shades of conflict in similar incidents, and it is not immediately clear where to draw the line between what satisfies an internal armed civil conflict and what is better classified as international terrorism or similar. For example, the low-level insurgency between the Egyptian government and al-Gamaa al-Islamiyya, which included the infamous 1997 Luxor attack by gunmen that killed over 50 foreign tourists, is included as a low-intensity conflict but the deaths of foreigners killed by related terrorist attacks do not appear to be included in the annual battle-related deaths for the conflict (Gerges 2000; Gleditsch et al. 2002b). For this instance, including foreign civilian deaths or not makes no difference since the number of local casualties exceeds 25 but not 999, making the case a low-intensity conflict regardless. (As described later, one of the advantages of binary coding for intensity is the ability to more comfortably handle ambiguous deaths or uncertain information.) There are other cases that are less clear: for example, n 1998, the Oromo Liberation Front allegedly attacked a Kenyan army base in Wajir, in northeast Kenya, killing 142 people (Global Terrorism Database 2010b), but the Oromo Liberation Front is primarily identified as a territorial conflict actor in Ethiopia, not Kenya, and even then not until 1999 (Gleditsch et al. 2002b). Given the research resources available, it is not known how many incidents might fall into similar disputed categories, but the number of ambiguous cases in Africa appears to be relatively limited less then 200 terrorism incidents between 1970 and 1999 in sub-Saharan Africa where the individual attack resulted in more than 25 deaths, according to the Global Terrorism Database (Global Terrorism Database 2010a). In a general review of these incidents, many of them appear to either already be included in the conflicts outlined in the ACD or to target non-governmental facilities.

In regards to the second concern – that parties may make false claims about their objectives – it is a fair concern but one that cannot be easily addressed without privileged information about a conflict actor's true intent. Moreover, there may be multiple points of view about the intent of the conflict within an insurgent organization or a government, so defining any single purpose is challenging on its own. While there may be a slight concern in this area, the ACD determination appears to be fair in all of the cases considered for my project.

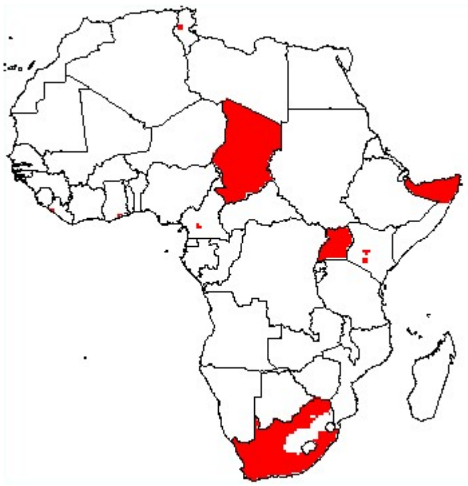
The final concern – that the nature of the incompatibility between the conflict actors can change over time – is a fair concern that is also acknowledged by the ACD authors. In discussing the conflict in Sudan between the Sudanese People's Liberation Movement (SPLM) and the government, the authors acknowledge that "by 1997 the incompatibility had widened to concern governmental power in addition to the territorial dispute" (Gleditsch et al. 2002b), but it continues to be marked as a governmental conflict in the database. The instance in it appears to be one of the few cases with such a clear transition in Africa, so while these circumstances of mixed incompatibilities raise minor concerns, it is not enough to invalidate the approach.

For their study, Buhaug and Rød limit their scope of conflict data to only internal conflicts – removing the interstate conflicts – and preserve the ACD distinction between incompatibilities for government and territory, and I have maintained that scope for this thesis project. Buhaug and Rød excluded internal armed conflicts where one of the actors was a colonial power, so Mozambique between 1964-1974 is not included, even though it resulted in over 3000 deaths per year for several years in the 1970s (Lacina and Gleditsch 2005, 2006; Buhaug and Rød 2006). In the study, Buhaug and Rød present two distinct maps – based on a single regression covering several decades – with one each for governmental and territorial conflicts (Buhaug and Rød 2006), but do not present a combined map at any point. For Buhaug and Rød, the distinction is preserved because "these types of conflict are likely to be shaped partly by different conditions," but without a common regression for both types of conflict, the differentiation is assumed before it is demonstrated (Buhaug and Rød 2006).



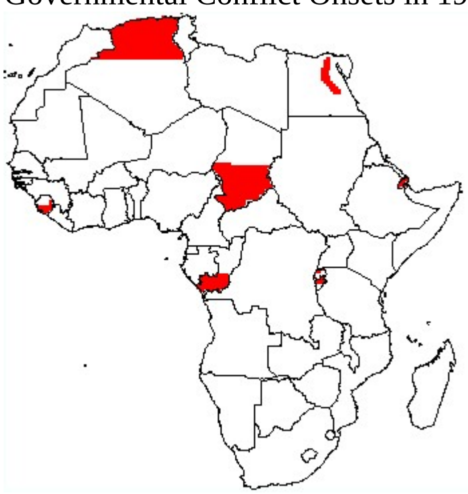
Governmental Conflict Onsets in 1970-1974

Governmental Conflict Onsets in 1974-1979



Governmental Conflict Onsets in 1980-1984

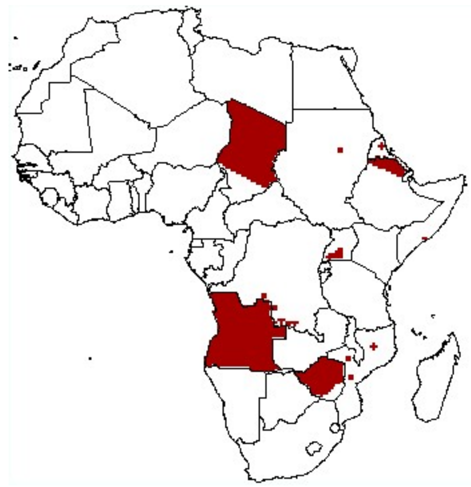
Governmental Conflict Onsets in 1985-1989



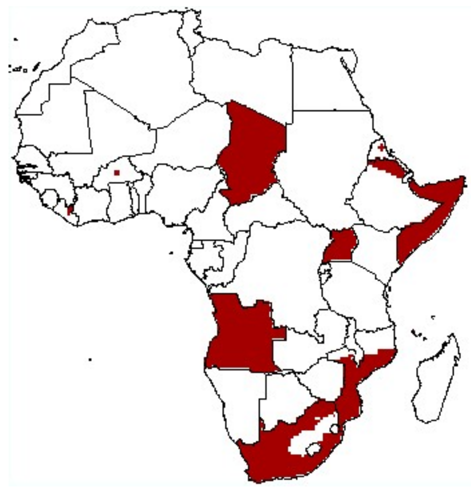
Governmental Conflict Onsets in 1990-1994

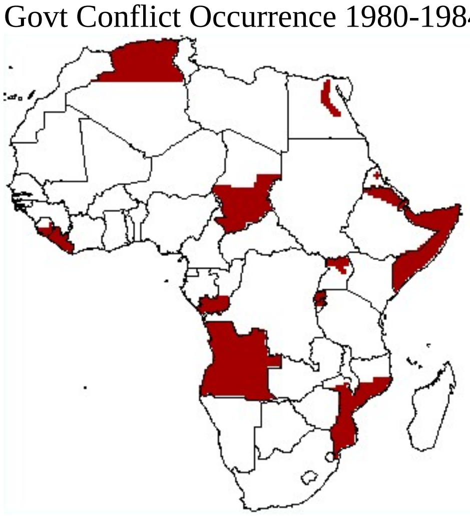
Governmental Conflict Onsets in 1995-1999

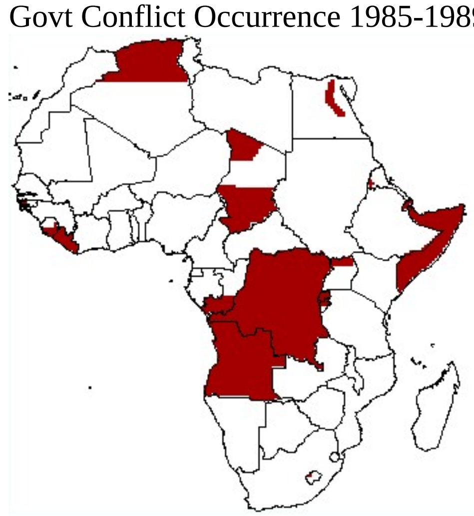

Govt Conflict Occurrence 1970-1974


Govt Conflict Occurrence 1975-1979

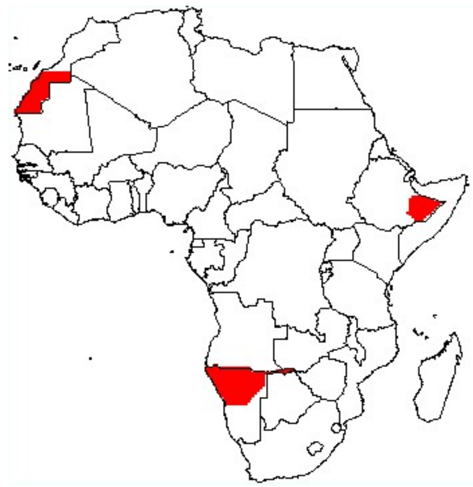

Govt Conflict Occurrence 1980-1984


Govt Conflict Occurrence 1985-1989


Govt Conflict Occurrence 1990-1994


Govt Conflict Occurrence 1990-1999

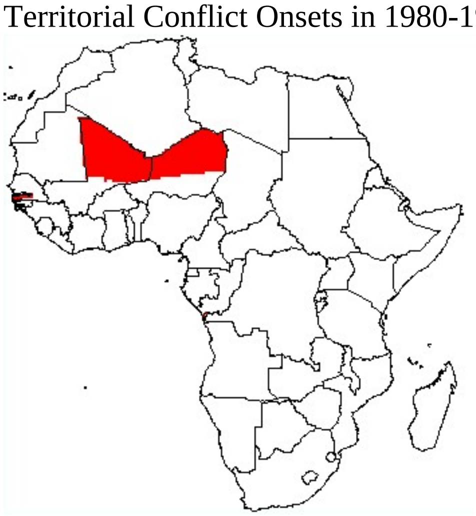
Territorial Conflict Onsets in 1970-1974[15]



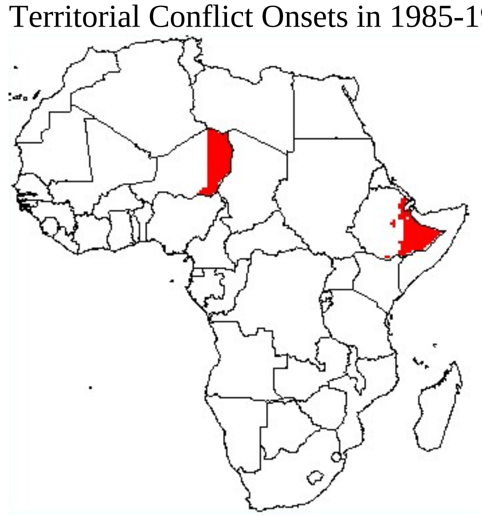
Territorial Conflict Onsets in 1974-1979

Territorial Conflict Onsets in 1980-1984

Territorial Conflict Onsets in 1985-1989



Territorial Conflict Onsets in 1990-1994



Territorial Conflict Onsets in 1995-1999

[15] There were no territorial conflict onsets between 1970-1974, which created problems with lack of variance in the variable.

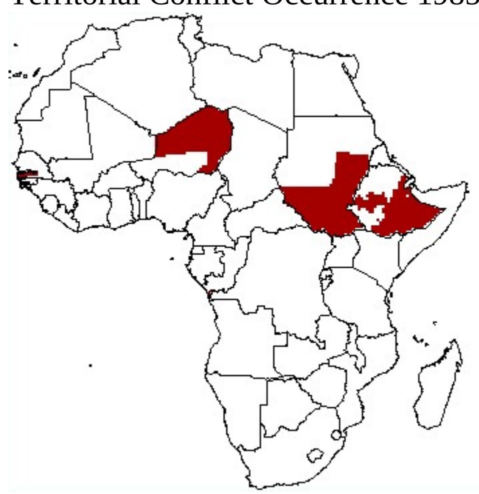

Territorial Conflict Occurrence 1970-1974


Territorial Conflict Occurrence 1975-1979


Territorial Conflict Occurrence 1980-1984


Territorial Conflict Occurrence 1985-1989


Territorial Conflict Occurrence 1990-1994


Territorial Conflict Occurrence 1990-1999

Conflict Data: Accounting for Intensity of Conflicts

Accounting for intensity is important, particularly because this study aims to assess a wide range of conflicts rather than just "wars" with more than 1000 battle-related deaths per year. Thus, the conflict data used integrates the polygons that have been designed by Rød and Buhaug with the number of battle-related deaths per year that are provided on a *conflict*-year basis in the latest version of the Uppsala/PRIO Armed Conflict dataset. Taken in five-year periods, it is possible to present a level of intensity for the conflict in the five-year period. Coding the intensity of a conflict as the aggregate number of battle-related deaths in a five-year period, according to the existing data in the Uppsala/PRIO Armed Conflict dataset, is still not a perfect fit. In a number of cases, such as Chad and the Democratic Republic of the Congo, multiple conflicts occur during the same time and in the same space. In these instances, the grid cells were coded for the most intense conflict present. In other words, a grid cell could conceivably have two low-intensity conflicts occurring – each resulting in 750 battle-deaths per year – so a total of 1500 battle-related deaths are occurring in the grid cell any given year, but it would still be classed as experiencing a low-intensity conflict. The reason for this coding – rather than taking the aggregate number of battle-related deaths in any grid cell – or across any period – is that does not represent the true nature of the conflicts occurring individually within the space of the grid cell.

Another major difficulty with mapping intensity is how to account for changes in the intensity of a conflict. For instance, how should one interpret the measurement of a situation that gradually evolves from a low-intensity conflict to a war?

Since I am using five-year periods, this problem only seriously presents itself in cases where the conflict develops in a war one or more periods after its initial onset. In the entire dataset, there were

My initial plan had been to use the Buhaug-Rød coding for intensity of the conflict, which generally followed the Cumulative Intensity coding in the Uppsala/PRIO dataset. Unfortunately, there appeared to be some inconsistency in the coding of the dataset

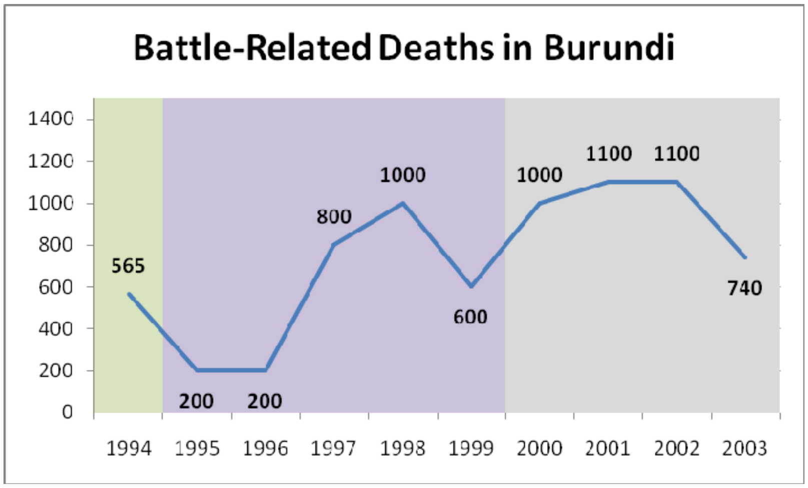


Figure 16. This graph of the battle-related deaths, drawn from PRIO data, demonstrates the challenges of identifying the onset of low-intensity conflict and civil war in conflicts whose intensities change over time. Based on a review of the data and the case, I opted to code a civil war onset in 1994, although there are periods where PRIO indicates less than 1000 deaths per year. Other scholars have code the conflict as starting as a civil war (not a minor conflict) in 1991 (Doyle and Sambanis 2000; Collier and Hoeffler 2004) or 1993 (Fearon and Laitin 2003b).

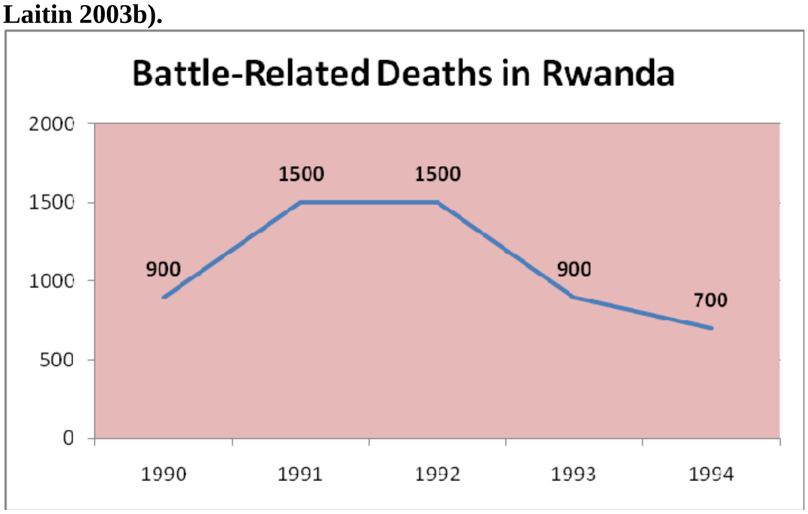


Figure 18. Battle-related deaths in Rwanda peaked in the same five-year period, easing coding questions.

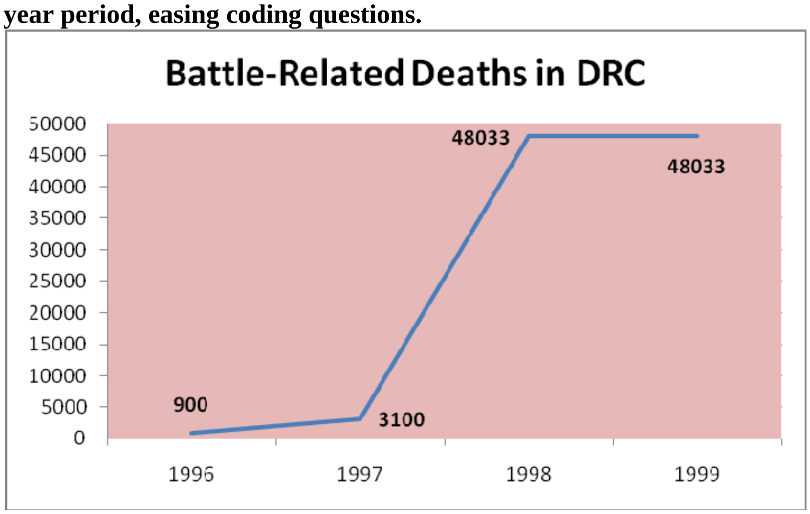
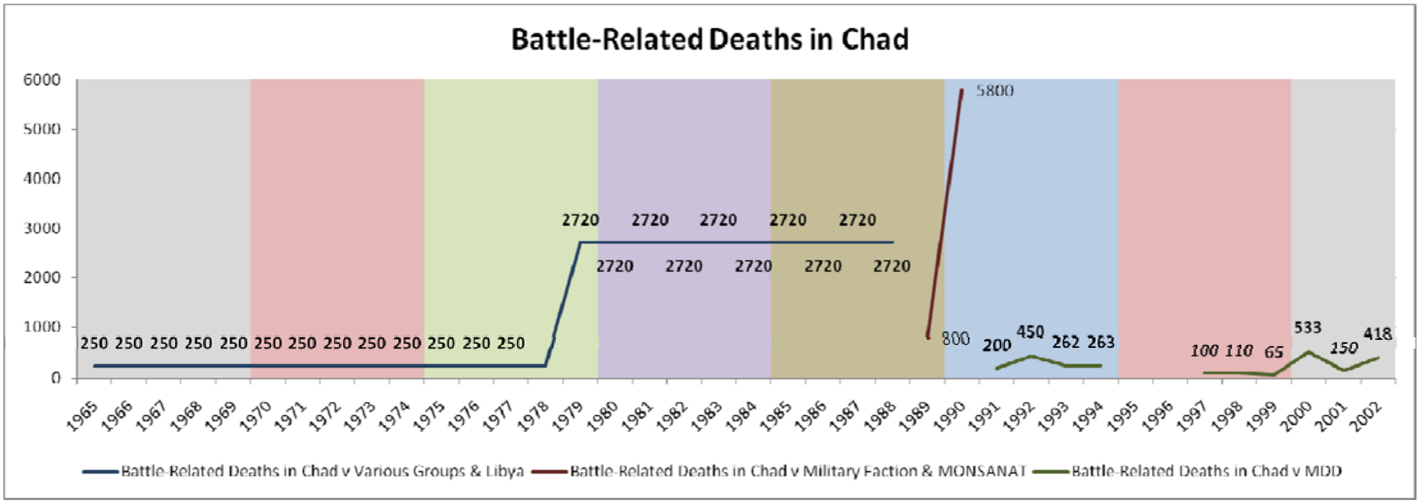


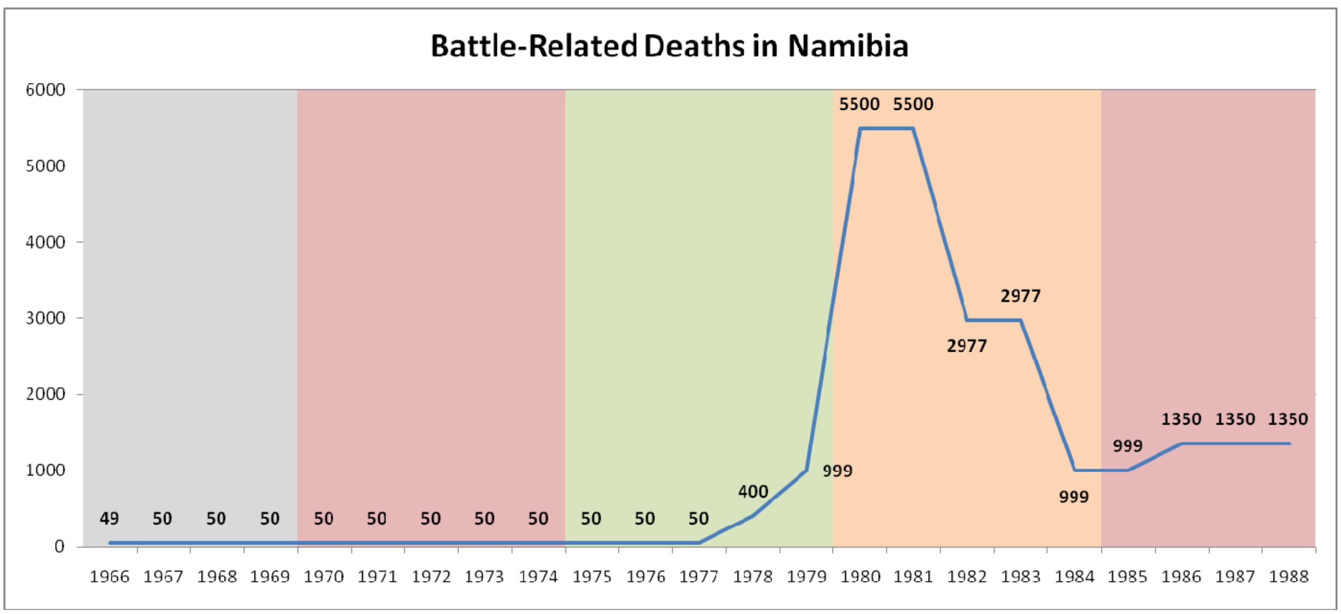
Figure 17. Battle-related deaths over time in the DRC show a marked increase from 1996, when it technically onset as a minor armed conflict before almost immediately becoming a civil war.

for intensity by Buhaug and Rød. While Chad's conflict, which starts in 1962 but does not achieve "civil war" status until 1974 is coded as a war (Intensity = 3) for its entirety, other conflicts, such in the Democratic Republic of the Congo in 1996, have their first year marked as a minor conflict (Intensity = 1), but then subsequent years are marked as "civil war" (Intensity = 3) when the battle-related deaths increase over 1000. In other words, some conflicts appear to be marked based on the highest intensity achieved while other conflicts appear to be marked based on the intensity of that particular year. Often, in these cases of confusion between the peak intensity of a conflict and the starting intensity, additional participants appear to have joined the conflict in the intervening years – such as the additional of Angola and Uganda as conflict actors in DRC in 1997. These conflicts still did not receive unique start dates, though, and were only differentiated by the arrival of new actors. To rectify this confusion between peak and starting intensity, I recoded the intensities of every individual conflict (as defined by having a unique start date) as the highest level of conflict achieved. So, DRC 1996 is marked as the onset of a war even though the particular battle-related deaths that year were only 900. Marking the onset in 1996 of the conflict in DRC as the beginning of a low-intensity conflict would seem not to fit well since the DRC developed into one of the worst wars in Africa, killing almost 50,000 people per year. These cases were handled on a case-by-case basis. For instance, in Buhaug-Rød dataset in Chad from 1965-1979, a "war" is marked as occurring even though only 250 people were dying per year, and, according to the Uppsala/PRIO battle deaths data, conflict does not increase to over 2000 battle-related deaths until1980 (Lacina and Gleditsch 2005, 2006). In this instance, I coded a war onset in 1980 in Chad since the dramatic ten-fold increase in violence marked a serious change in the conflict and so that the coding of the Chadian conflict would be more comparable to other studies that mark a war onset in that time period (Collier and Hoeffler 2004). Chad was the most difficult case to code by far, and it has been for many others working on similar projects. As Sambanis noted, "there is very high disagreement in coding the war [in Chad] among major datasets"

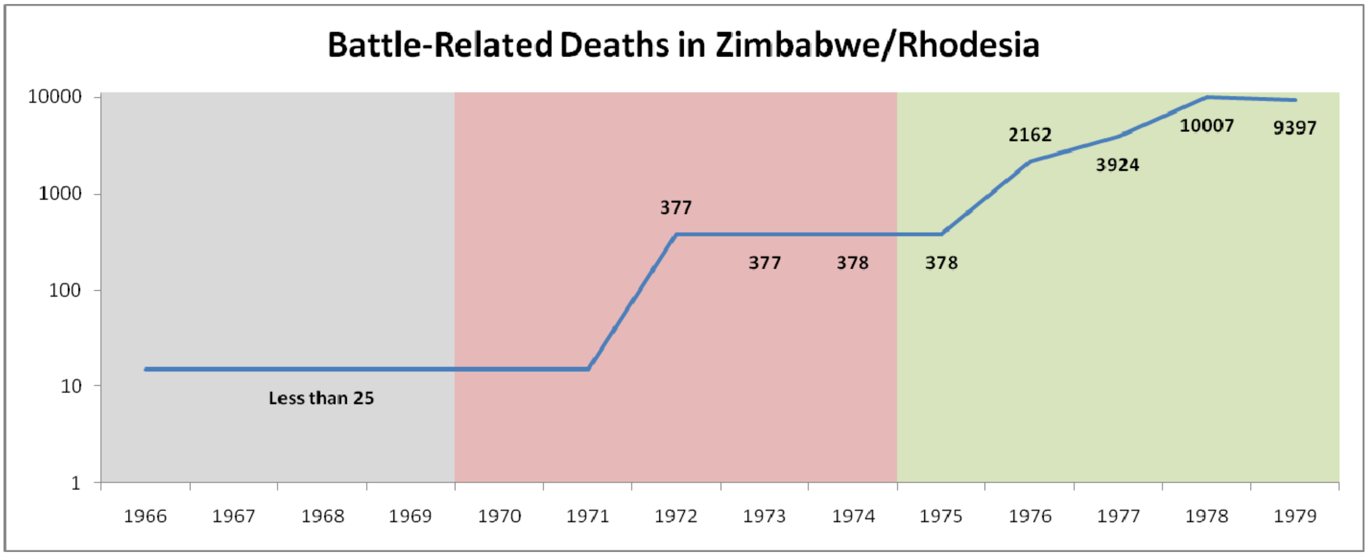


(Sambanis 2004), and different datasets code anywhere from one minor armed conflict from 1996-2001 (Gleditsch et al. 2002b), one war from 1980-1988 (Singer and Small 1994) to four wars from 1965-72, 1978-86, 1983-86, and 1989-94 (Regan 1996), among other variations. Based on a brief case review, events in 1979, including the decision by one faction of the government, led by the Northern Prime Minister Hissein Habre, to send an army to fight the national army in the capital, N'Djamena, suggests that 1979 marks a sufficient transition in the conflict to mark an onset in 1980 when a brief Organization of African Unity-sponsored ceasefire broke down immediately (Sambanis 2004). I coded all previous years as an active minor civil conflict since while Buhaug and Rød mark the entirety of the period as having "various groups" new conflict actors appeared to take a role in 1979/1980 justifying a new onset. I also coded the renewal of the conflict in 1989 as a war onset even though it only caused 800 battle-related deaths that year because the subsequent year caused over 5000 (Lacina and Gleditsch 2006). Last, I did not code a new onset in 1997, despite the break in battle-deaths in the Uppsala/PRIO dataset, because the conflict did not have a new start date or new conflict actors.

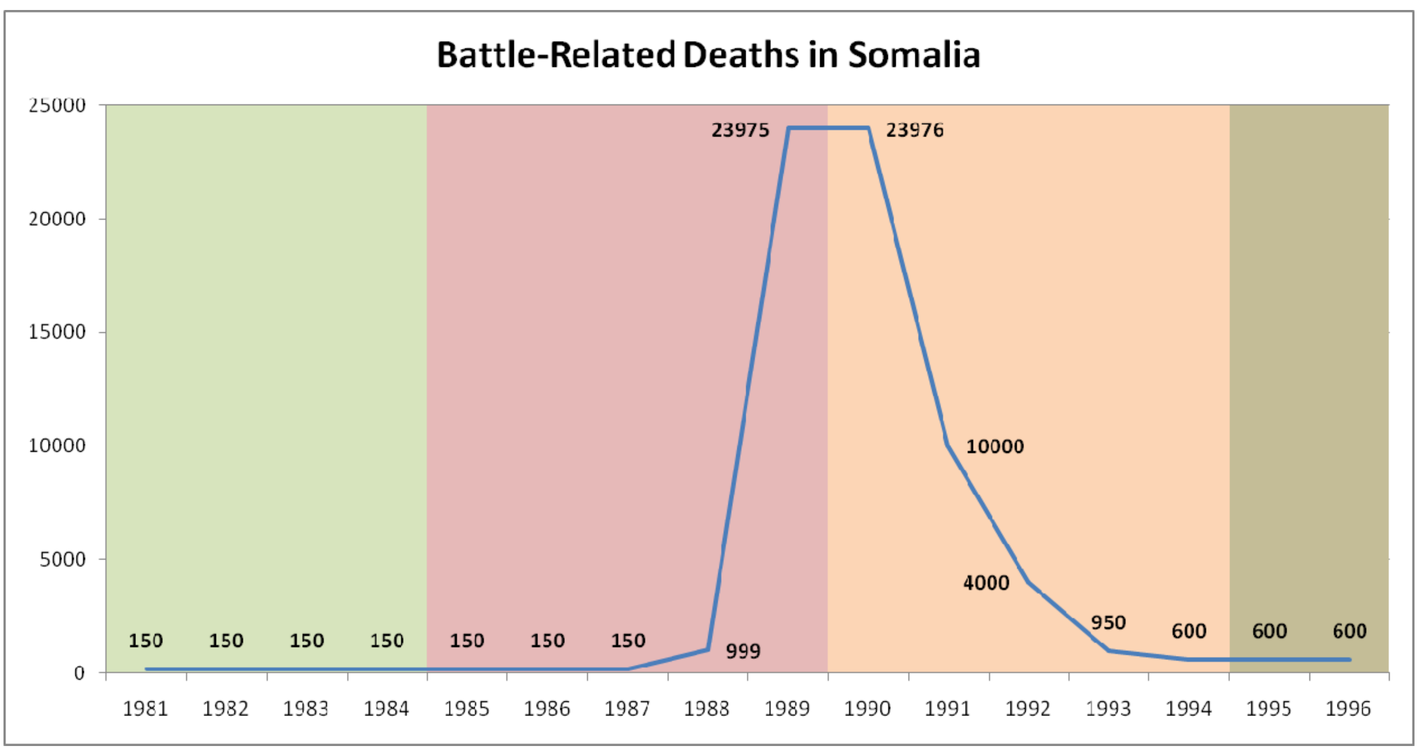
In a comparable case, the war that separated Namibia and South Africa started as a minor civil conflict in 1966 and continued to have approximately 50 battle-related deaths per year from 1966 to 1978, when it began a precipitous climb to 5500 by 1980. As a result, I have coded a civil war onset in Namibia in 1978. Similar to Chad, coding conventions on Namibia vary: neither Fearon and Laitin nor Collier and Hoeffler code any war whatsoever (Fearon and Laitin 2003b, 2003a; Collier and Hoeffler 2004), while Sambanis allows that "it is possible that the sustained violence and effective resistance criteria were not met until after 1977 (so one could code a low-level violence before 1977 and civil war in 1977)" (Sambanis 2004). Again, I have based my decision on the duration of the conflict, the available data on battle-related deaths, comparable datasets in aggregate and disaggregate studies, and a brief case study review.



In a similar case in Zimbabwe/Rhodesia 1972 and 1976, spikes in violence appears to have changes an armed conflict started in 1966. In this case, I chose to code a civil war onset at 1972 since it marked the beginning of an almost uninterrupted trend of upward



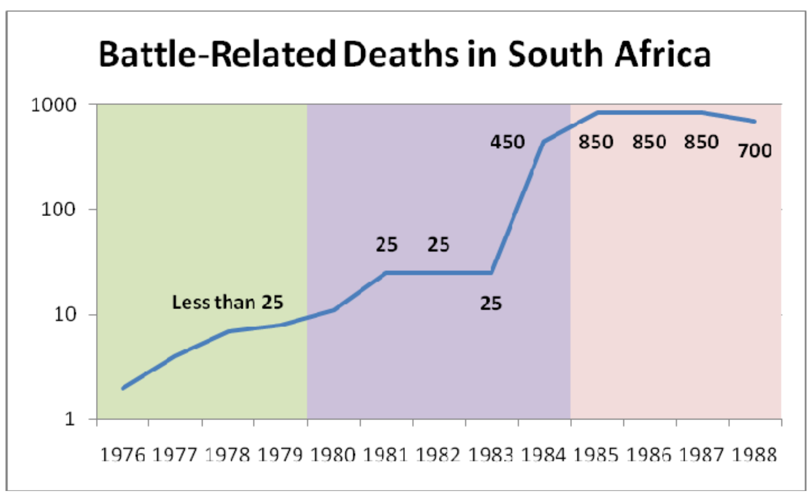
violence towards war, and since that coding would be more consistent with other comparable studies (Collier and Hoeffler 2004; Sambanis 2001, 2004)..



Coding for Somalia was also a challenge. In Somalia, there appear to have be several periods of conflict – a minor armed conflict with less than 1000 battle-related deaths between 1981-1988, a civil war between 1988-1991, and then an ongoing period of conflict, which is not well measured by the Uppsala/PRIO dataset, possibly because of the difficulty of designating any conflict actor as "the government." While Fearon and Laitin mark the beginning of the Somali civil war as 1981, research by Sambanis as well as Collier and Hoeffler suggest a later start date near 1988 (Fearon and Laitin 2003b; Collier and Hoeffler 2004; Sambanis 2004). Of note, the Correlates of War project does not code a civil war during this period (Sambanis 2004). I believe that the most reasonable method for parsing these conflict periods is to mark a minor civil conflict from 1981-1988, and a civil war onset starting in 1988 that does not conclude by the end of the dataset in 1999. Coding the period from 1991 onwards as a separate conflict could make sense because the Barre regime was overthrown in that year, but it is not clear that enough of the conflict actors or fundamental situation changed, despite a brief ceasefire, to suggest a separate coding for subsequent conflict.

In Liberia, the civil war is marked as starting in 1989, consistent with several other sources (Fearon and Laitin 2003a; Sambanis 2001; Collier and Hoeffler 2004), even though the Uppsala/PRIO dataset only has 100 battle-related deaths for 1989 (and 5000 for the year following) (Lacina and Gleditsch 2005, 2006). Similarly in Sierra Leone, although the Uppsala/PRIO battle-related deaths database does not mark over 1000 battle-related deaths until 1998, I have coded the start of the civil war as occurring in 1991, consistent with several other studies (Sambanis 2001; Fearon and Laitin 2003b; Collier and Hoeffler 2004), and I have not coded a new war onset in 1997, although others do (Sambanis 2001; Collier and Hoeffler 2004). In another similar case in Algeria, I assign the onset of the conflict in 1991, even though the Uppsala/PRIO dataset has only 100 battle-related deaths to be consistent with other studies (Fearon and Laitin 2003b; Collier and Hoeffler 2004) and since no separate minor armed conflict preceded the war. The battle-related deaths at the end of 1991 were merely the beginning of the much larger war that followed. Last, in Congro-Brazzaville, I assigned two onsets – one for minor armed conflict in 1991 between the forces loyal to Pascal Lissouba and those loyal to Denis Sassou-Nguesso that ended in 1997 and the other onset for a civil war that started in 1997 between different factions of the Cobra and Ninja militias, loyal to Sassou-Nguesso and a former mayor of Brazzaville, Kolelas, respectively (Sambanis 2004). Given that the dataset in my study is grouped into five-year periods, the primary purpose of this reassignment is to indicate a civil war onset rather than a low-intensity conflict onset.

Note: In the original Uppsala/PRIO dataset for battle-related deaths, there appears to be an error in the cumulative intensity (CumInt) of Ethiopia's conflict with Eritrea from 1962-1991.

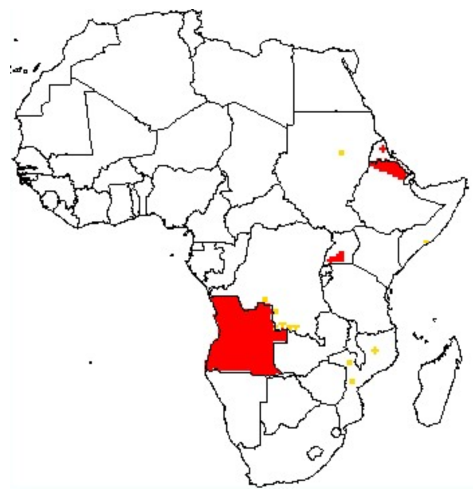


I spent considerable time and energy considering how to properly handle the transition from conflicts that were low-intensity to conflicts that became wars, so what consideration, if any, should be given to any battle-related deaths that occurred before the conflict reached 25 battle-related deaths per year? Ultimately, it is not possible to provide marking for the beginning of a conflict but before it reaches the level of 25 battle-related deaths. Although the Uppsala/PRIO dataset does provide information on original start date of a conflict and the first time that the conflict reaches 25 battle-related deaths/year – see South Africa for an example of the difference – it is not possible to include this information in the coding or regression analysis. For every study, there must be a minimum threshold of conflict, and while my thesis lowers that threshold to 25 battle-related deaths per year, it is important to maintain some minimum to avoid introducing too much noise into the analysis.

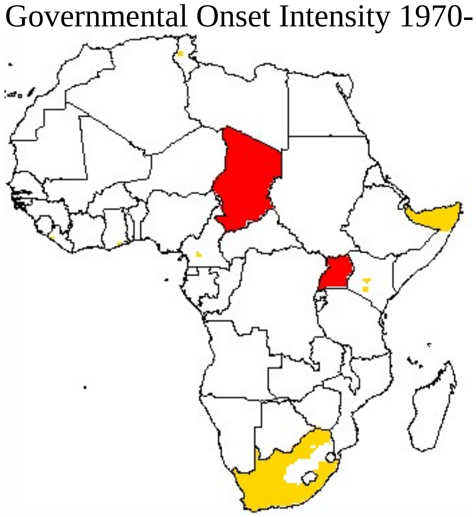
Another difficulty with the variable for intensity of conflict is that it does not account for the "density" of the conflict. In other words, because the available data only indicates how many deaths occurred in a conflict, not where they occurred within the conflict space, it is difficult to know which areas were most impacted. Thus, the number of deaths for the conflict-year has been applied to all cells in that conflict-year's polygon. In a case such as the Democratic Republic of the Congo, where conflict is indicated to have existed throughout the country, this approach undoubtedly leads to under-reporting casualties in some areas (such as the hard-hit eastern part of the country), and over-reporting in other areas that were less affected by the civil war.

While there are some techniques that might produce a more reasonable intensity map, such as weighting the number of battle deaths by the population density in each cell or dividing the total number of deaths evenly over the total number of grid cells for that conflict-year, each of these approaches has its flaws as well. For instance, weighting conflict cells by their population density would assume that battle deaths are highest in urban areas, which is not always the case, particularly in rural-driven or "sons of the soil" rebellions (Fearon and Laitin 2009). The best way to track conflicts, and their intensity, would be through the method used by ACLED, which marks the location and details of each battle in a conflict. However, this level of precision in data collection was not feasible for a study of the entire continent of Africa, although we can hope that such refined data may someday be widely available (Raleigh and Hegre 2005). As a result, the best way to view the findings from the intensity regression analysis is that certain factors lead to a given cell being part of conflict(s) that caused *x* number of battle-related deaths.

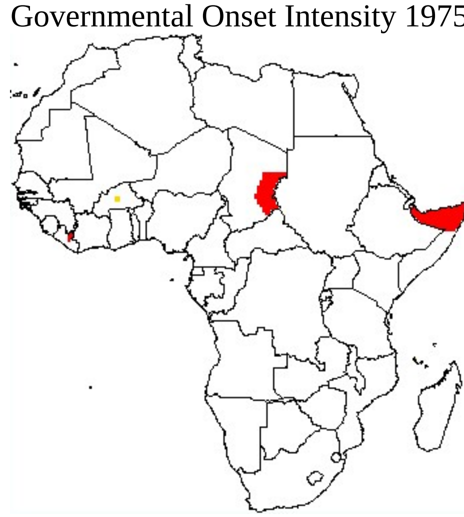
Governmental Onset Intensity 1970-1974



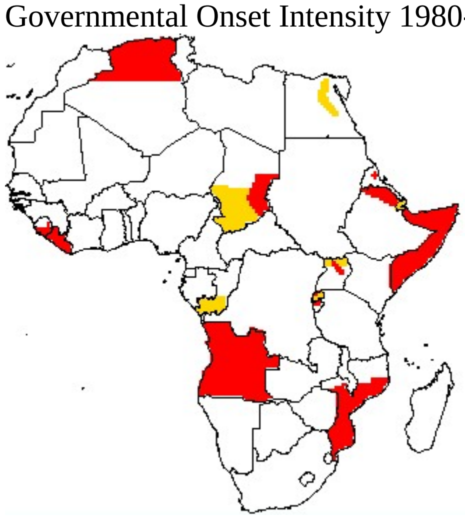
Governmental Onset Intensity 1975-1979



Governmental Onset Intensity 1980-1984

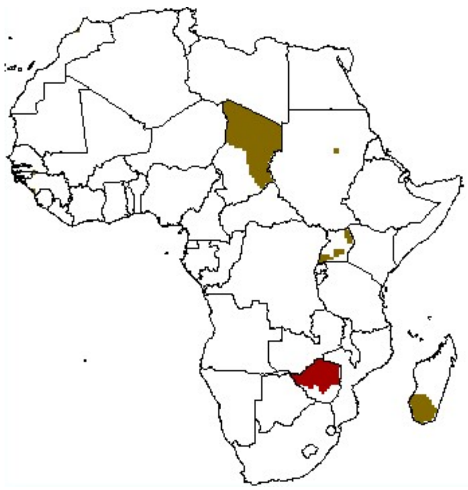


Governmental Onset Intensity 1985-1989

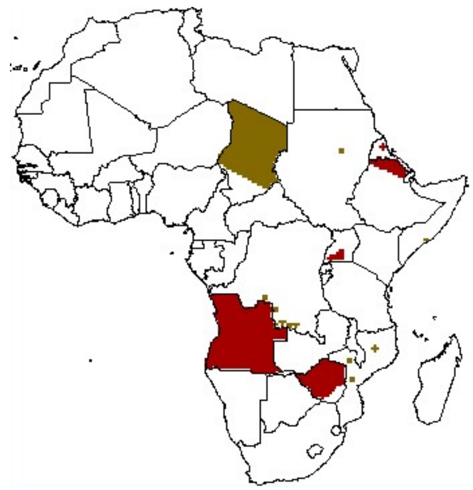


Governmental Onset Intensity 1990-1994

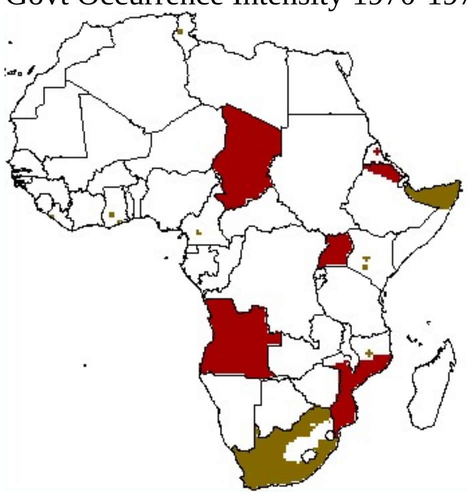
Governmental Onset Intensity 1995-1999



Govt Occurrence Intensity 1970-1974

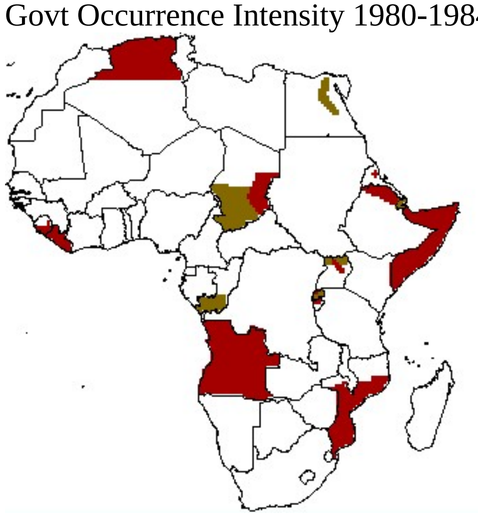


Govt Occurrence Intensity 1975-1979

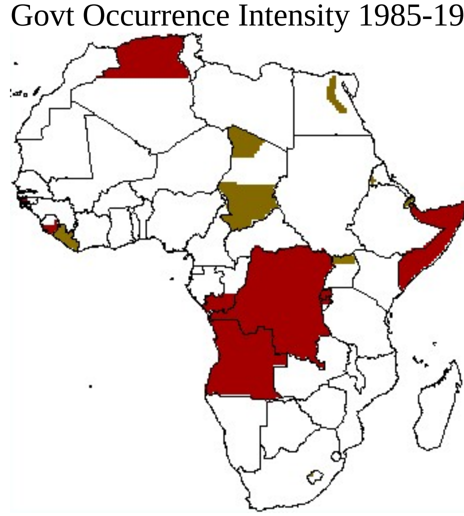


Govt Occurrence Intensity 1980-1984

Govt Occurrence Intensity 1985-1989



Govt Occurrence Intensity 1990-1994



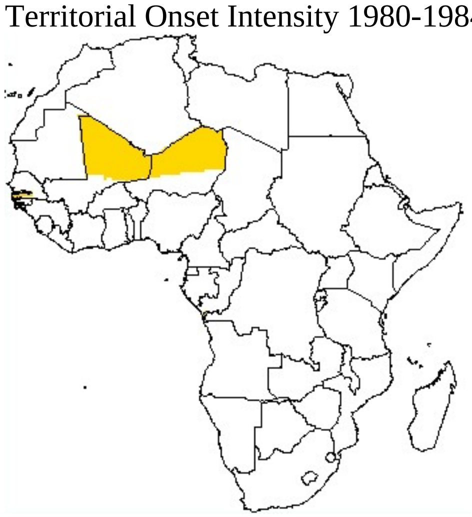
Govt Occurrence Intensity 1995-1999

Territorial Onset Intensity 1970-1974

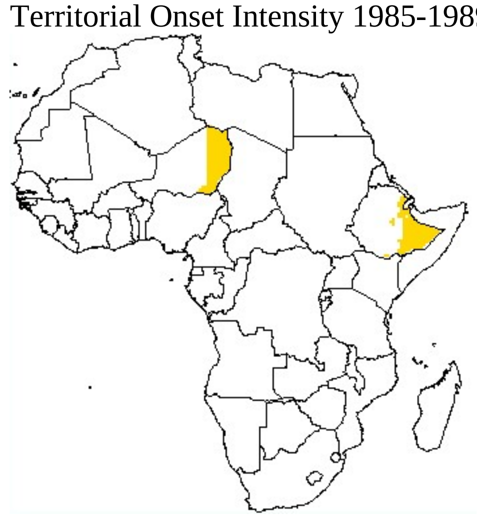
Territorial Onset Intensity 1975-1979

Territorial Onset Intensity 1980-1984

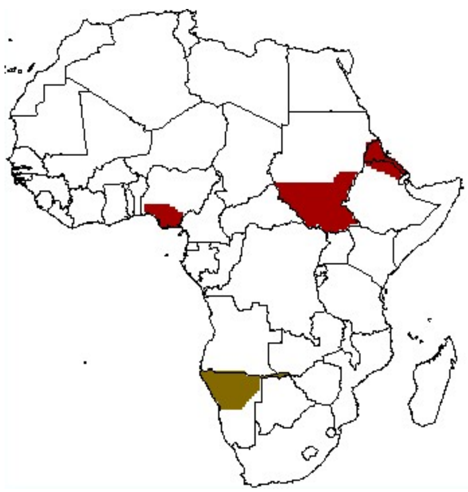
Territorial Onset Intensity 1985-1989



Territorial Onset Intensity 1990-1994

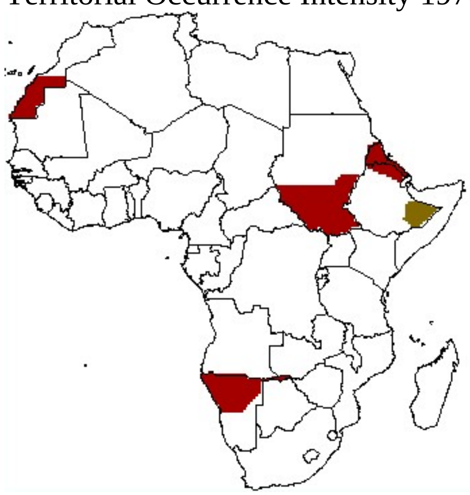


Territorial Onset Intensity 1995-1999



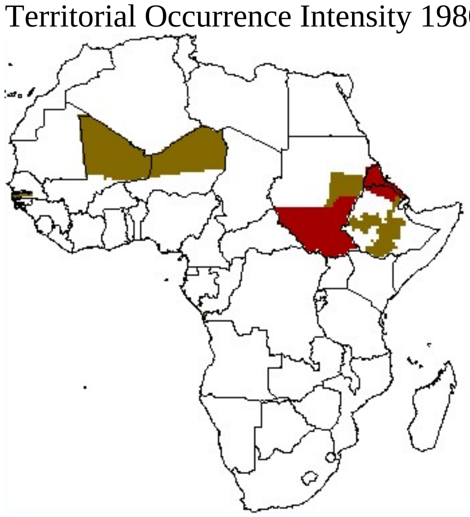
Territorial Occurrence Intensity 1970-1974

Territorial Occurrence Intensity 1975-1979

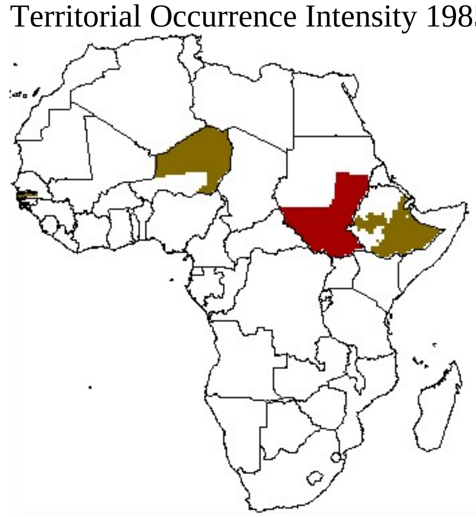


Territorial Occurrence Intensity 1980-1984

Territorial Occurrence Intensity 1985-1989



Territorial Occurrence Intensity 1990-1994



Territorial Occurrence Intensity 1995-1999

Spatial Lag

The spatial lag variable – or the correction for spatial autocorrelation – is perhaps the most significant variable in any regression of disaggregate data for conflict. As Rød and Buhaug demonstrate, the method by which autocorrelation is taken into account through spatial lag variables can affect the findings on civil war in a meaningful way (Rød and Buhaug 2007). Unsurprisingly, the most significant factor regarding whether a given cell is likely to experience conflict is whether the cell next to it is experiencing conflict, so the design of the spatial lag variable – which accounts for conflict in nearby grid cells – is important to the entire regression outcome. For this project, the coding of the spatial lag variable differs significantly from Buhaug and Rød's approach. Buhaug and Rød code their spatial lag as a state-level number for the percentage of conflict experienced in the state over the years and cells included. (This process is described in greater depth in the replication of one of their models with GWR in Annex 1.) The result is a single value for spatial lag for all of the cells in a state. This approach seems to run counter to the overall aim of the disaggregate study to assess units smaller than the state. By assessing spatial lag at the aggregate level, the lag variable in the Buhaug-Rød study assumes that conflict is a disaggregate variable when it occurs but an aggregate variable when it is lagged. It also provides the same spatial lag correction for grid cells – regardless of whether the cells are located immediately next to a conflict or far away from a conflicts. For conflicts in large states that experience conflicts in only a portion of their territory, such as Sudan, Mali or Niger, – a statewide spatial lag variable presents an unusual picture of the impact of the conflict.

For my thesis, I chose to use a kernel-based density is calculated for the conflicts around each grid cell, instead of generating state-level percentages as a variable for spatial lag. This approach counts the number of cells either with conflict onsets, occurring conflict, or battle deaths (depending on which independent variable the spatial lag is targeting) and then sums them based on a quadratic function. This calculation was performed in ArcGIS 9.1, which uses a quadratic equation specified by Silverman in his analysis of kernel density analyses (Silverman 1986).[16]

Kernel densities call for defining the size of the kernel – effectively the radius around which each cell will be assessed. While increasing the size of an already large kernel will not have much impact (because the additional cells assessed will have very low weights), enlarging smaller kernels can have a substantial impact on their results. As kernel radius increases, the density formed around the underlying shape decreases. So, for small radii, the kernel density can closely mirror the underlying shape, but for large radii, the kernel density much less closely mirrors the underlying conflict. This situation is laid out in the three radii options shown, with a 2.5 degree kernel radius on the far left, a 5.0 degree kernel radius in the middle, and a 10.0 kernel radius on the far right. (These maps are the spatial lag data for grid cells where conflict is occurring with more than 25 battle-related deaths per year in the 1990-1994 period.)

[16] For those seeking to reproduce this spatial lag, the specific expression is equation 4.5 on page 76 of his book, but ArcGIS will automatically draw on this equation when producing kernel densities.

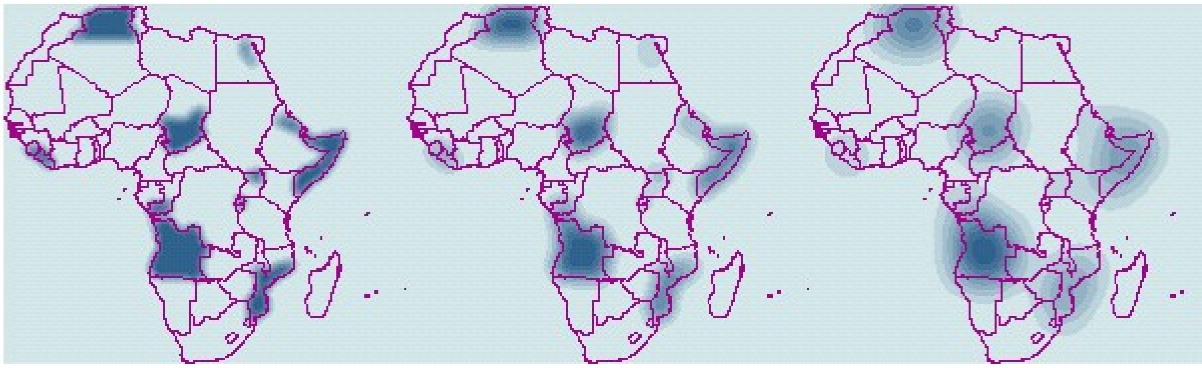


Figure 19. The images above show the impact of the kernel density on the spatial lag variable. On the right, the kernel density is set to 2.5 degrees. In the middle, the kernel density is set to 5.0 degrees. On the left the kernel density is set to 10.0 degrees. The underlying conflict data is all conflicts from 1990-1994.

Choosing an appropriate radius for the kernels is not a precise science – any choice will lead to a somewhat arbitrary breakpoint where the spatial lag ends and becomes completely zero. It is tempting to pick a very large radius and thus reduce the number of cells without spatial lag effects and potentially reduce the significance of the spatial lag in the regression results – thereby potentially increasing the significance of the other variables. Selecting a very small radius that would ensure that the spatial lag more closely matches the curves of each conflict polygon, but that would make the spatial lag closer to the value of the dependent variable itself, risking making all other factors insignificant. For this project, a radius of 5 degrees has been chosen, not because it represents the difficult-to-define "ideal" spatial lag distance but because it represents a middle option and no common practice could be determined from past literature. One could easily select another kernel distance or perhaps generate a formula for optimizing the spatial lag in some manner, but for the purposes of my project, a mid-range value should serve well.

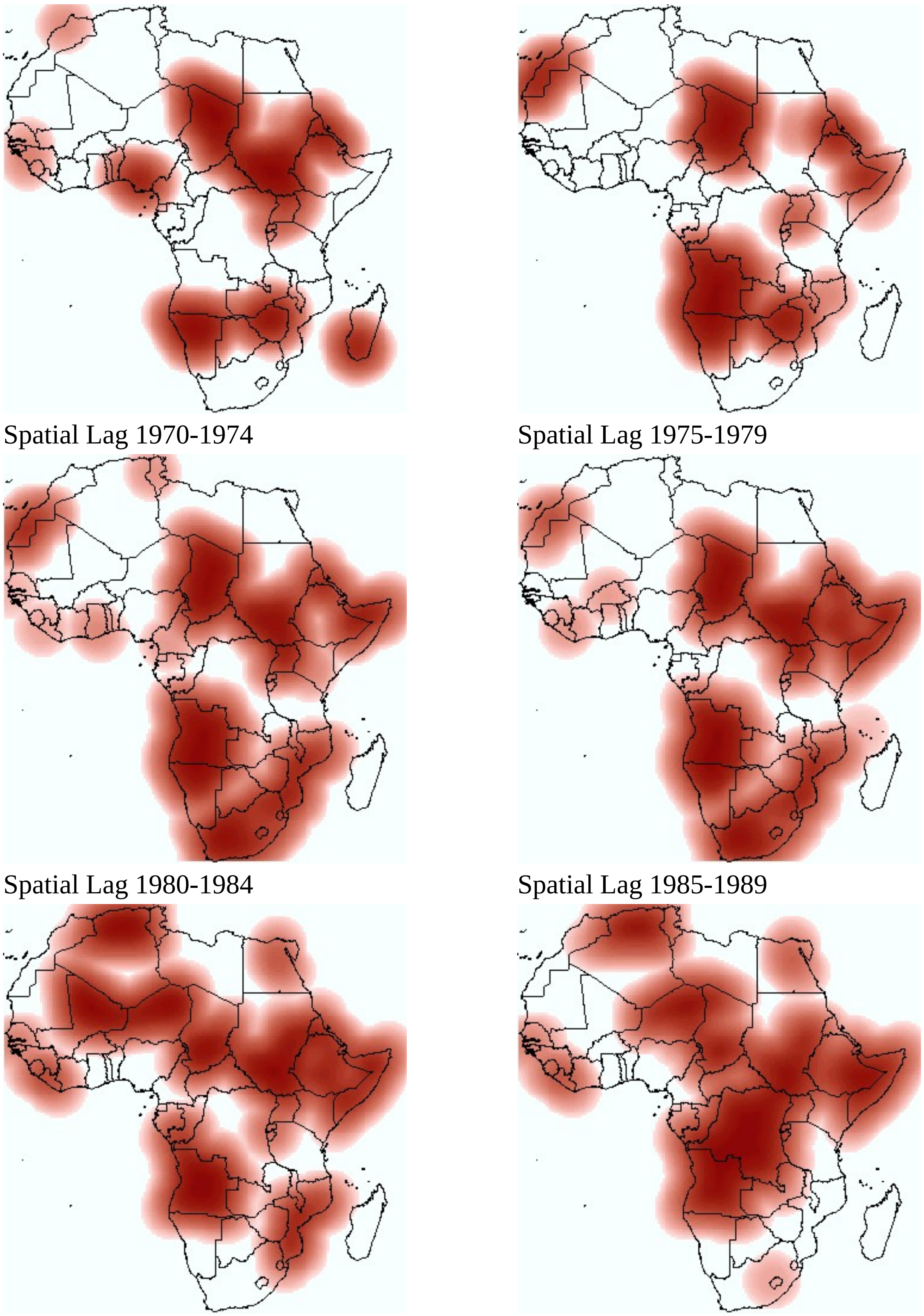
While there will likely always be some elements of a spatial lag variable that raise concerns, it is important to highlight any possible distortions that it may cause in the regression results. One concerning element of the density function, and anything over a very small radius, is that border regions in general appear to have a much lower spatial lag. Even though many of the conflicts are coded as being equally present in the center of the country as in the borders, the spatial lag variable is always stronger in the geographic center of a conflict because it encompasses more conflict cells in the density kernel. On the edges of a conflict, which often correspond with the borders of a country, the spatial lag variable is smaller because more of the cells included in the density kernel come from outside the country and the conflict. In the results section, the findings on the significance of borders were generally quite robust, but a small caveat should be attached to them because of the method by which I generated the spatial lag variable.

Alternatives to the kernel density approach were considered, but I opted not to pursue them for this project. I examined using a straight-line distance for the spatial lag variable, but I rejected it on three major conceptual grounds. First, it would come very close to putting the dependent variable on both sides of the regression equation. In other words, since the distance to a conflict from a conflict cell is 0, the spatial lag independent variable would almost perfectly model the exact dependent variable. This situation would massively reduce the significance of any other factor in the regression. Second, a straight-line distance spatial lag would make it difficult to model situations in which there are several nearby conflicts. From my perspective, a grid cell that is 100 km from one conflict with no other conflicts for at least 2000 km should have a smaller spatial lag than a grid cell that is 100 km from one conflict and 150 km from another. And, the latter grid cell should have a smaller spatial lag than a grid cell that has two conflicts, each 100 km away, on either side. With simple straight line directions, all conflicts other than the closest one would have been ignored. Last, the simple straight line distance would have made it difficult to take into account the differing levels of conflict in each grid cell. The ability to model differing weights is irrelevant in modeling the onset or occurrence of conflict because both are binary variables, but choosing a kernel density approach gave me greater flexibility in my approach to intensity.

Hegre and Raleigh used an alternate method of measuring spatial lag, which I also considered. For their study, they tracked the distance between any conflict event and the nearest other conflict event (Hegre and Raleigh 2005; Raleigh and Hegre 2009). While this metric was well-suited to their data, which measured events as short as a day, it would not have been appropriate for the structure of the conflict data used in my study because my study uses multiple conflict polygons, sometimes overlapping, rather than individual conflict points. Thus, measuring the distance from one conflict to another was not a suitable approach for my study, and kernel densities appeared more reasonable.

Once I chose the method for my spatial lag – kernel density – I needed to choose the conflict variable from which it would be generated. In specific, should territorial and governmental conflicts receive separate spatial lag variables? If I separated the spatial lag by conflict type, then the variable would have assumed that territorial conflict cannot be a contributor to governmental conflict, or vice versa. Some previous work suggests that separating spatial lag by conflict type may be worthwhile: initial findings by Buhaug and Gleditsch indicate that secessionist conflicts cause “contagion” and spread neighboring areas, but that conflicts motivated by attempts to control of the central government are not similarly contagious (Buhaug and Gleditsch 2005). As I reviewed the specific cases involved, differentiating the conflicts by the aims of their actors did not make sense. For example, separating territorial and governmental conflicts would imply that the impact of the onset governmental war in Angola in the late 1970s did not impact the onset of territorial war in northern Namibia (then South Africa) in the same period. In fact, the cases appear linked: the same rebels – the National Union for the Total Independence of Angola (UNITA) – who were fighting in Angola were also present as conflict actors in Namibia during the Southwest African People's Organization (SWAPO) liberation movement (Arnold 1995; Sambanis 2004). Moreover, the two accords that brought both wars to a close were negotiated and signed at similar times because the conflict actors were intertwined (Berridge 1989). Similar cases exist in the conflict linkages between territorial conflict in eastern Ethiopia and the governmental conflict in Somalia, where flows of refugees, links between ethnic groups, and the impact of conflict on development tied the two conflicts closely (Lewis 1989; Clark 1993; Gundel 2002). Based on this case information, I opted include all conflicts regardless of the aims of the actors in the spatial lag variable.

Having consolidated conflicts by the aims of the actors for the spatial lag, I also considered whether to consolidate the conflicts based on their level of intensity. In other words, is it relevant in constructing a spatial lag variable that some conflicts are low-intensity and others are civil wars? Based on a review of the cases and the theory, I decided that the situation was not sufficiently clear to warrant separating the conflicts by intensity. In some cases, such as the series of low-intensity conflicts and civil wars in Chad over the 1980s and 1990s, the conflicts are linked with common conflict actors and an evolving political situation that led to continued conflict (Sambanis 2004). A similar case exists in the conflicts between the DRC, Uganda, Rwanda, and Burundi in which common ethnic groups and even the governments themselves have played a role in neighboring conflicts regardless of their intensity (Parker et al. 2000; Wallensteen and Sollenberg 2001; Sambanis 2004; Olsson and Fors 2004). Thus, separating conflicts by their intensity for the purpose of creating the spatial lag did not appear justified either. So, for the spatial lag variable took all conflicts occurring – regardless of intent or intensity – and used them to create a kernel density spatial lag variable for each five year period.



Spatial Lag 1970-1974
Spatial Lag 1975-1979
Spatial Lag 1980-1984
Spatial Lag 1985-1989
Spatial Lag 1990-1994
Spatial Lag 1995-1999

Figure 20. These maps represent the spatial lag for conflicts occurring in Africa in the six five-year periods studied. The spatial lags were generated with kernel densities set to a five-degree radius.

Time Period Lag

Because this analysis will be one of the first conflict analyses to combine aggregate political and economic data with disaggregate geographic data, it will also be necessary to include a time period lag for conflict in the previous time period. The time period lag accounts for whether a grid cell is experiencing conflict largely because it experienced conflict in the previous period. For this variable, there were two approaches considered: first, the variable could be a standard direct lag of the occurrence of conflict in the grid cell in the previous, or second, the variable could be a lag of the spatial lag variable. The second option is appealing because it takes into account conflict that may be nearby in the previous period but not necessarily in the grid cell itself. While this option was appealing, I opted to implement the first option – a direct lag of the occurrence of conflict in the previous period for use in this analysis. A direct lag option the advantage of being more appropriate for a dynamic logit analysis where it is necessary to make a clean, binary distinction between the onset year and subsequent years. If the spatial lag variable had been lagged a period to serve as the temporal lag variable as well, there would not have been a clean distinction because of the values of the kernel density function are not binary. Considering the alternatives, I decided that preserving the temporal lag variable as a simple, binary measure of whether conflict occurred in the grid cell in the previous period was a more reasonable approach.

Conclusion

The project aims to make contributions to the academic literature by integrating existing data sources and methodologies in a hybrid disaggregate approach with aggregate variables, and to apply this hybrid approach to the entire continent of Africa. By taking the two major sources of data – temporally varying political and economic data from the national aggregate level and spatially varying geographic data from the local disaggregate level – and integrating them to provide results that have not yet been presented at any major paper or source. The results from this approach then help to inform the theoretical debate over what factors are most commonly associated with civil conflicts.

Incorporating the latest datasets, such as from the conflict polygons designed by Rød and the ethnic group mapping from the Geo-Referencing Ethnic Groups (GREG) project, provides the most up-to-date metrics on key variables and should make the results relevant to the present state of civil war literature. The data sources described above are largely already in existence – the central contribution of the study is not new datasets – but specification of the data used in this study, particularly in the area of Polity IV scores, terrain, forests, and distance to capitals/borders, has been refined so that the specification more closely matches the underlying theoretical hypothesis. On the independent variable side, accounting for temporal changes in governance and economics will improve on past disaggregate studies, and, on the dependent variable side, including information on the intensity of conflict will provide a more refined understanding of what factors are most significantly correlated with lower or higher intensity conflict. Moreover, the data used also combines variables from various major theoretical causes of civil war: using data on ethnic groups, political and economic variables, and geographic and structural variables.

Examining these factors together at a disaggregate level has not been conducted before for all of Africa.

The statistical results from the data analysis should help to add to a debate within and between disaggregate and aggregate studies. It should help to shed light on how local and national factors may interact in determining the onset, occurrence, and intensity of civil war. The theoretical contributions are unlikely to debunk any existing arguments in the aggregate or disaggregate fields, but the contributions are more likely to explain why current disagreements between the two may be reconcilable or not through integrated analysis of national and local factors and noting distinctions between different types and intensities of civil conflict.

Table 2. The table below summarizes the data sources that were used to construct the dataset for my thesis. Transformations were applied to many of these original dataset to derive new specifications, but these dataset are the source for the underlying raw data.

Type of Data	**Source**
Battle-related deaths	Uppsala/PRIO Armed Conflict dataset
Conflict polygons	Buhaug-Rød conflict polygon shapefiles
Diamonds	DIADATA (PRIO)
Elevation data	US Geological Survey GTOPO30 projections
Ethnic configuration borders	Geo-Referencing Ethnic Groups (GREG) based on *Atlas Naradov Mira*
Forestry data	Scores calculated based on images from UNEP-GRID satellite images circa 1995 from US Geological Survey
Geographic data (proximity to capitals and borders)	Calculated based on rasterized digital map
Oil	PETRODATA (PRIO)
Other non-fuel minerals	US Geological Survey
Political data	Polity IV scores
Population	UNEP-GRID geo-referenced data
Rainfall data	Interpolated from US National Weather Service 50-Year Reconstruction Project
State economic data	Greed and Grievance dataset (Collier and Hoeffler) UNU WIDER dataset
State population	Aggregated from UNEP-GRID data National Material Capabilities dataset (COW)

CHAPTER 4: REVIEW OF HYPOTHESES ON CORRELATES OF CONFLICT

The range of causes considered for civil war has been wide and varied over the course of its study, but often the causes considered have aimed to explain all civil wars considered with a single regression approach. In a limited number of cases, civil wars have been differentiated by their intent, such as territorial or governmental aims, or their alleged cause, such as "ethnic" wars, or when empires and colonies fought. One of the challenges in comparing these various findings is that they focus on such a broad spectrum of conflict that some variables take on an ambiguous significance. The aim of this section is to provide a short synopsis of the major findings on some of the most commonly considered variables. This section will point out where findings have been conclusive, leading to reliable hypotheses about the models tested in this project, and where findings have been less clear, inviting more theoretical speculation and empirical testing of hypotheses. This section also aims to clarify some of the specific instances in which variables are expected to be significant – such as only territorial conflict or only low-intensity conflict.

While the following list of hypotheses might appear individually narrow in scope and without a central focus, the robustness of the dataset assembled invites an opportunity to test a range of variables that have been the subject of previous study without clear results. Moreover, the hybrid approach taken by this project offers a rare opportunity to examine political, economic, geographic, and environmental variables together in a disaggregated approach. In some cases, geographic factors that have been heavily considered in past disaggregate studies – such as distance to a state capital – have rarely been considered alongside the political system of the state. Similarly, some economic variables – such as GDP per capita or statewide growth of GDP – have not been included in disaggregate studies in the past, suggesting that a checking their relevance in this context would be productive. Many of the hypotheses in the following section are not new, but they are new to this approach or are being tested with new controls.

Last, this section provides a catalogued method for assessing whether the broad approach of distinguishing between territorial and governmental conflicts, as well as between low-intensity conflicts and civil wars, is relevant for the study of civil war. It may be, as some authors have implicitly assumed, that all or almost all civil wars should be explained with a single analysis, often a regression approach. If there are no serious distinctions between the territorial and governmental conflicts, or between conflicts of different intensities, then this approach may be warranted. However, if there are significant distinctions – which must be assessed on a variable-by-variable basis – then distinguishing between different aims of conflicts or intensities of conflict may be a more productive approach to describing the correlates of civil conflict.

Gross Domestic Product (GDP) and Economic Performance

Gross domestic product is one of the most commonly used variables, but it is often left open to many interpretations for what it represents. Gross domestic product per capita can be used to reflect state power and national development, but the specific aspects of "state power" being operationalized are not always clear. As Collier and Hoeffler point out, "so many characteristics are correlated with per capita income that…the proxy is open to other interpretations" (Collier and Hoeffler 2004). Fearon and Laitin find GDP per capita to be "strongly significant in both a statistical and substantive sense" but the variable is used to represent range of factors: per capita income proxies for "police and counterinsurgent weakness," "a state's overall financial, administrative, police, and military capabilities," "more developed countries with terrain more 'disciplined' by roads and rural society penetrated by central administration," opportunity cost to recruits in a civil war, "state military and police strength relative to insurgents," and just generally "state strength" (Fearon and Laitin 2003b; Haring-Smith 2005). Regardless of its varied explanations, GDP is consistently the most robust predictor of conflict in aggregate studies. In a wide range of studies low GDP has corresponded with high civil war risk (Fearon and Laitin 2003b; Collier and Hoeffler 2004; Hegre and Sambanis 2006). Thus, for the purpose of this study,

Hypothesis 1. *Low country GDP corresponds with higher local conflict risk.*

Hypothesis 2. *Low GDP per capita corresponds with higher local conflict risk.*

Similarly,

Hypothesis 3. *Low country GDP growth corresponds with higher local conflict risk.*

As noted earlier, GDP measures in aggregate studies do not reflect the distribution of income or development across the territory of a state, though, so the finding that low GDP correlates with high conflict is only half the issue.

It is noteworthy that there are a number of case studies of conflicts that have been associated with strong economic performance. For example, in September 2008, violence erupted in Santa Cruz, Bolivia – an eastern province that benefits from natural gas deposits and has experienced strong economic growth - when the Bolivian government in La Paz announced plans to re-distribute some of the natural gas wealth ("Fresh violence in Bolivia stokes civil war fears" 2008). Similar cases of violence from autonomous regions with stronger economic performance can be seen in Somaliland in northern Somalia and northern Italy (Menkhaus 2007; Giordano 2000). It would be ideal if these local trends in GDP could be analyzed through this study but it is unlikely that GDP itself will be observable because no easily useable dataset of local GDP over the time period covered exists. Instead, some independent proxies for economic growth can be considered, as described below.

One of the major difficulties of using GDP as an independent variable in models of violent civil conflict is that it is not necessarily independent. GDP – the economic performance of a county – is affected by civil conflicts and may also be affected by the tensions that lead to violence. In other words, the assumption of the models presented is that because the GDP per capita is low then the state becomes more prone to violent conflict. However, if the GDP is decreasing because the tensions leading to conflict – hostile relations between groups, initial roadblocks by rebel groups, large non-violent protests, countrywide labor strikes – are also slowing the economy then the low GDP may be caused by the coming conflict, rather than the other way around. This potentially endogenous relationship between GDP and conflict makes it desirable to find a measure of economic strength that operates independently of violent conflict. Miguel et. al attempt to use data on rainfall as a measure for economic growth since rainfall is independent of conflict patterns and strongly affects economic performance is heavily agricultural countries in Africa (Miguel 2004). Their results suggest that rainfall may have a strong relationship with conflict, but they were called into question by Jensen and Gleditsch who argued that including countries that participated in other civil wars – such as former Zaire participating in neighboring conflicts – skewed the results considerably (Jensen 2008). Jensen and Gleditsch also argue for using disaggregate statistics on rainfall and conflict rather than the national statistics used by Miguel et. al. (While Miguel et al. draw their data from disaggregate sources of gridded rainfall data, they aggregate the information in order to make it comparable with the available civil war data. With the Buhaug-Rød conflict polygons, this aggregation should not be necessary.)

In examining the rainfall data, it becomes clear that there is little change in climatic patterns between regions. Unsurprisingly, the equatorial regions consistently get more rain than the Saharan regions. Given the varied economic performances across Africa but the relatively consistent rainfall patterns when viewed across periods of years, there is reason to question whether there is a modern connection between absolute levels of rainfall and economic growth. There are two more potentially relevant variable specifications – first, the difference in rainfall from the previous period, which could be measured in percentage or absolute difference in inches; and second, the difference per capita. Since rainfall is could proxy for the amount of economic growth or available resources, both significant consideration is how much development is occurring per capita, these local measures of rainfall still merit consideration. Conversely, since some of the theory that suggests rainfall is important is concerned with resource shortages, it is important to consider how many people are trying to divide up the available rainfall. Indeed, a sharp spike in population growth combined with a marked decrease in rainfall would seem to present a particularly troubling situation. Thus,

Hypothesis 4. *Increases in rainfall overall correspond with reduced risk of conflict.*

Hypothesis 5. *Increases in rainfall per capita correspond with reduced risk of conflict.*

These various hypotheses on rainfall also test the basic significance of local economic or resource-availability factors for economic growth. Since aggregate studies are limited in their ability to assess local economic conditions with a country, taking the disaggregate approach opens a window into the relevance of these factors.

Horizontal Inequality (Income Distribution)

There have been theories about civil conflict since the days of Aristotle, and even Aristotle performed rudimentary tests on his theories. Aristotle wrote, “Inferiors revolt in order that they may be equal, and equals that they may be superior. Such is the state of mind which creates revolutions” (Aristotle, Politics, Book V, Section II). This basic theory – that conflicts are caused by what would later be termed “horizontal inequality” – produces a testable hypothesis. Aristotle tests it merely by assessing the range of city-states that he knows about. For example, Aristotle comments that a horizontal inequality resulted "at Argos, where the Argives, after their army had been cut to pieces on the seventh day of the month by Cleomenes the Lacedaemonian, were compelled to admit to citizen some of their Perioeci"(Aristotle, Politics, Book V, Section III). Perioeci were inhabitants but not citizens in Sparta and lived concentrated on the geographic and social outskirts of the territory. According to Aristotle, the admission of this out-group into the Argives' society resulted in a horizontal inequality between Argives and Perioeci that led to civil war.

One might make a similar argument that many modern cases, such as the series of conflicts in the Balkans or Central Africa, was spurred by the effect of horizontal inequalities, situations where groups of people who are poorer or wealthier than neighboring groups come into conflict. Modern studies of horizontal inequalities have also used case studies. Stewart assessed nine case studies to determine that horizontal inequalities caused significant political disturbances (Stewart 2002), although his case selection – Chiapas in Mexico, Indians in Fiji, the Acholi in Uganda, Bumiputera in Malaysia, Tamils in Sri Lanka, blacks in South Africa, Catholics in Northern Ireland, blacks in the United States, and blacks in Brazil – could be viewed as heavily focused on selecting cases of known tension or civil conflict, if not civil war.

In empirical studies, the horizontal inequalities tested by Collier and Hoeffler use data on income and land inequalities. While their data is limited in the number of countries that it covers (the Gini coefficients were some of the most sparse data in their dataset), they also do not find statistical significance from the aggregate factors (Collier and Hoeffler 2004). The Gini coefficients are not ideal for testing the concept of horizontal inequality because they reflect only the overall distribution of wealth or land in a country, not the relative standing in income or land between groups, which is part of the core of the concept.

These centuries-old theories of civil war can be tested using more advanced statistical methods. For example, Østby presents data from 55 developing countries from 1986-2003, collected through *Demographic and Health Surveys*, that reveals strong statistical evidence, through regressions, that inequalities between the capital and outlying regions are strongly connected to the onset of conflict (Østby 2007). In country-specific studies, the results on the impact of horizontal inequality on the location of conflict have been mixed. In Nepal, assessments of horizontal inequality in land-holding and income among the 75 districts revealed that deaths from that civil conflict concentrated in districts with high rates of landlessness or significantly less income that the capital, Kathmandu (Murshed and Gates 2003). In Liberia, a similar study that assessed where conflict deaths occurred relative to the income of the province found that richer provinces had greater levels of conflict deaths (Hegre et al. 2007). The likely explanation for the difference between the two results is that the insurgency in Nepal was weak relative to the government, meaning that most battles took place near poorer rebel bases, whereas the insurgency in Liberia was strong, pushing battles into richer government areas. Methodologically, the transition from the study on Nepal to the study on Liberia reflects a improvement in focus of the analysis; in Nepal, data was based on district-level information then regressed independently of each other, whereas in the Liberia study, data was plotted on a grid covering Liberia in 76km^2 cells and were regressed against geo-referenced data. While this methodological difference does not make one set of results more valid than the other since the studies were focused on different countries and different conflicts, it does reflect the growing sophistication of civil conflict research. This form of disaggregated research focused on income distribution has not been broadly expanded beyond single country studies.

In testing inequality, this study draws on data from the aggregate and disaggregate level. Using the measures for land inequality and income inequality from the Collier-Hoeffler dataset, this study assesses what impact overall state-level inequality has for the risk of conflict in any location. Unfortunately, this data does not exist for many of the countries in question, which eliminates much of the data from any regression analysis. For measures at the disaggregate level, there are no exact economic data on the sub-region, as discussed earlier. As a replacement, this project takes the data from the annual rainfall to determine which regions have received more rainfall compared to other regions in a given country. Thus, there is one general hypothesis at the aggregate level and two hypotheses at the disaggregate level:

Hypothesis 6. Countries with higher rates of income inequality experience a higher risk of conflict onset and occurrence throughout their territory.

Hypothesis 7. Countries with higher rates of land inequality experience a higher risk of conflict onset and occurrence throughout their territory.

At the disaggregate level,

Hypothesis 8. Areas with more than the state average for rainfall have a greater risk of governmental conflict.

And,

Hypothesis 9. Areas with less than the state average for rainfall have a greater risk for territorial conflict.

The hypotheses for governmental conflict and territorial conflict are separated in the case of rainfall at the disaggregate level because it seems more likely that underinvestment in poorer, drier locations would lead to a territorial conflict driven by the lack of resources. In contrast, in better developed areas that enjoyed more than the state average for rainfall, the economic development that comes with rainfall suggests that capturing the state itself would be more attractive. It could also put the wetter areas in a stronger position relative the strength of the state, creating a larger opportunity for insurgents from wetter areas to capture the state government. Last, distinctions in rainfall within a country could delineate different cultural groups as sustain differences in rainfall patterns could lead to distinctions in cultural practice. In this case, regrettably, it is difficult to directly connect a single operational instrument to greater risk of conflict, but it does provide an opportunity to examine broadly whether local differences in the distribution of resources – rainfall – can have an impact on conflict risk using a variable that is completely exogenous from conflict risk, unlike many other resource variables.

Rough Terrain

Rough terrain has been one of the variables where the greatest theoretical consensus exists. From Machiavelli to Mao Zedong, mountains and forests are described as places of refuge for those who are not prepared to fight (Zedong 1937). As Machiavelli wrote, “When a Captain wants to fight almost secure in not being routed, he arranges his army in a place where he has a safe refuge nearby, either amid swamps or mountains or in a powerful city; for, in this manner, he cannot be pursued by the enemy, but the enemy cannot be pursued by him” (Machiavelli 1520). Machiavelli also argues that wooded and mountainous places are best for ambushes, in line with guerilla warfare, saying in Book 4 of the *Art of War*, “in wooded or mountainous places, and hence capable of ambush, part of your forces can be well hidden, so that the enemy may be assaulted, suddenly and without his expecting it, which will always be the cause of giving you the victory” (Machiavelli 1520). A rough examination of modern conflicts, from the Kurds in northern Iraq and the Hill Tracts of Bangladesh, the FARC of Colombia and Abu Sayyef of the Philipines, also makes the point in numerous examples of rebels who rely on hills, mountains, foliage, and forests to gain safe havens (Alvarez 2003; Rustad et al. 2008). These case studies, though, raise the question among academics of whether we grant too much credit to a number of high profile cases – such as the infamous Tora Bora mountains of Afghanistan – for creating conditions that are not actually significantly favorable to conflict on the whole. After all, on the other side, Switzerland has mountains but no major history of conflict. And one could argue, as Sun Tzu does at one point in the *Art of War*, that mountains are dangerous for military forces because they often lack sufficient access to food and water (Tzu). Thus, a number of statistical studies that focus on the relationship between rough terrain and civil conflict have been engaged (Buhaug and Gates 2002; Buhaug et al. 2005; Rustad et al. 2006;

Hendrix 2006). The discussion of rough terrain will have separate discussions of hills and mountains, which are mostly considered to provide only safe havens and constraining the reach of the state, and forests, which have also been connected to providing natural resources that can help to fund an insurgency, such as in the Philippines or Indonesia (Global Witness 2002; Rustad et al. 2008).

Hills and Mountains

Within early theoretical work, most theorists saw hills and mountains as safer territory, so long as food and water were available, and therefore a potential safe haven for rebels who needed a defensive territorial advantage to avoid direct military confrontations and compensate for inferior military forces. These conclusions were largely drawn by military theorists who were more concerned with fighting and winning wars than with the original causes of them (Machiavelli 1520; Zedong 1937). Scholars more concerned with the origins of war than necessarily how to best fight it have seen hills and mountains as a refuge for rebels as well as a limitation on the powers of the state, broadly conceived (Fearon and Laitin 2003b). There have been some analyses that have suggested that rough terrain can also have indirect effects, such as reducing the tax/GDP ratio, that can also increase the risk of civil war (Hendrix 2006). Indeed, there has also been the argument presented that the presence of mountains increases the number of minority groups who feel disconnected from the central government and may therefore seek separation.

The mechanism for believing that rough terrain increases conflict is often that rugged land provides terrain suitable for rebels and difficult to penetrate by government forces. However, it is possible that the impact occurs more centrally, which could explain the failure to find significant disaggregate results. Hendrix suggests that the ruggedness of the terrain actually reduces the capacity of the state – specifically the tax/GDP ratio – and therefore weakens the central state in all areas, making the whole state more prone to civil conflict (Hendrix 2006). (Unfortunately, reliable tax/GDP ratio data is so sparse in Africa that it cannot serve as a measure of state strength for this model.)

The relationship between mountains and armed conflict has also been the topic of significant study for international organizations. A 2002 statement by the Food and Agriculture Organization of the United Nations (FAO) noted, "Armed conflicts have emerged as serious threats to mountain ecosystems and their inhabitants with 23 of the 27 major, armed conflicts in the world in the year 1999 being fought in mountain regions" (FAO 2002). FAO Director-General Jacques Diouf also stated that "Mountain areas are home to most of the armed conflicts in the world as well as many of the world's poorest and least food-secure populations" (FAO 2002). This statement provides further suggestion that mountains should be considered both as a source of rebel refuge but also as a place that often develops disconnected and destitute minorities.

The combination of theoretical and case study evidence made rough terrain a common variable in country-year studies that were pursued. Fearon and Laitin returned a positive finding that the percent of mountainous territory in a country was linked to increased levels of civil war (Fearon and Laitin 2003b). Collier and Hoeffler also find that mountainous terrain is positively associated with civil war but found only a marginal level of significance in their "opportunity model." In the grievance and combined opportunity-grievance models, there is less clear evidence of the significance of mountainous terrain. Collier and Hoeffler's final conclusion, that "mountainous terrain might advantage rebels…[or] are correlated with unmeasured grievances" is notably weaker than Fearon and Laitin's findings (Collier and Hoeffler 2004). Siding with Fearon, Hegre and Sambanis find that rough terrain is important to the onset of civil war (Hegre and Sambanis 2006). Thus, based on aggregate results,

Hypothesis 10. *Higher elevation corresponds with higher conflict risk.*

Hypothesis 11. *Changes in elevation correspond with higher conflict risk.*

Hypothesis 12. *High slope gradients correspond with higher conflict risk.*

In disaggregate studies, though, there is significant evidence that actual fighting tends not to occur in areas where mountains exist. Rather, conflict polygons were associated with less mountainous areas of the states studied (Buhaug and Lujala 2005; Buhaug and Rød 2006). These studies, though, focused primarily on the areas of conflict, and researchers have suggested that mountainous (or forested) areas may provide a useful refuge for rebel bases and camps, even if they do not fight near them.

One reason for the disparate findings in aggregate and disaggregate studies may be that they tend to measure "rough terrain" using a measure of mountainous terrain, either as designated by an expert (Fearon and Laitin 2003b; Collier and Hoeffler 2004) or as provided by UNEP (Buhaug and Rød 2006). These findings focus on the percentage of mountainous terrain – the most significant form of roughness – but it may be that this terrain is *too rough* – that the costs associated with fighting or maintaining operations in the mountains exceeds the benefits even for rebels. Indeed, it may be that while mountains are challenging, hills are attractive. Thus,

Hypothesis 13. Hills – mid-level ranges for the elevation, change in elevation, and slope gradient – correspond with higher risks of conflict in a parabolic relationship, while very mountainous and very flat areas are less conflict-prone.

This distinction between mid-level and high-level roughness has not been made to any significant extent. With GIS technology and satellite imaging, though, it is possible to more reasonably identify these areas of mid-level roughness. Also, beyond terrain designation, it is not immediately clear what specifically constituted a "mountainous" area in previous studies when experts were used to designate "mountainous" areas of countries (Fearon and Laitin 2003b). Thus, for this study, a measure of the slope of each grid cell is used, presumably with higher slopes corresponding with rougher terrain. This data should create a continuous measure of roughness rather than simply a measure of the amount of mountains. A second measure could also be considered that would make the mean values for slope – i.e. the hills – the highest while making the very flat and very steep areas take lower values. This transformation might help highlight if the mid-level slopes are partially responsible for the differing results.

Foliage and Forests

The role of forests in conflict has had contradictory evidence and theory. On the one hand, there are the understandable theories that forests provide refuge for rebels and insurgents from the government forces. The impenetrability of the large groupings of trees provides shelter from both large mechanized incursions by government forces and provides some resources – such as plants and water – for sustaining a basic existence. On the level of country-year studies, Collier and Hoeffler could not find any statistically significant relationship and Fearon and Laitin did not include forestry data (Collier and Hoeffler 2004; Fearon and Laitin 2003b). In a comprehensive study focused on the relationship between forests and conflict, Rustad also found few significant relationships between forest cover and conflict, and what relationships were discovered were largely negative (Rustad 2005a; Rustad et al. 2006). Rustad suggests that the reason for the non-finding may be that there are either difficulties in the data or case studies have biased the common understanding of forests to view them as related to conflict. (One significant finding that Rustad did receive was that forest degradation increased conflict (Rustad 2005b), but testing that theory would require a more refined dataset of forest cover than is used in this study.) In 2005, the Food and Agriculture Organization of the United Nations published a report noting that there were several armed conflicts in forested regions, although it could not definitively show that forests were in fact a causal factor in these conflicts (FAO 2005). In disaggregate study, Buhaug and Rød also find that forests have little relationship to conflict, and, if any, it is likely a negative relationship with territorial conflict and a positive relationship with governmental conflict (Buhaug and Rød 2006).

Another point of view might argue that fighting in forests is difficult – for the government and the rebels, and not always close to key strategic "prizes" – whether it be a government capital or simply the major city of a secessionist target. Alternatively, nations with fewer forests may have fewer natural resources, which given the noted correlation between low GDP and civil war could lead to greater conflict in those nations. In research similar to this issue, Hauge and Ellingsen find that deforestation – the reduction in the amount of forest – leads to conflict (Hauge and Ellingsen 2001). This finding indicates that the process of reducing the amount of forest leads to conflict, but it may also stand to reason that the less forest that there is, the more conflict there will be as resources become scarce. Under these theories, conflicts might be fought in less densely forested territory.

Although there is not clear evidence in the empirical literature that we would expect forests to be significant, forestry data needs to be included, given the persistence of the belief that conflicts are aided by forests – and the anecdotal evidence from cases such as Laurent Nkunda's Tutsi rebels fighting in eastern Congo or the Lord's Resistance Army seeking refuge in forest reserves in northern Uganda and northeastern Congo (Prunier 2004). From the findings so far, it seems possible that the effect of forests is two-fold: that very dense forests increase conflict risk by providing refuge while very sparse forests increase conflict risk by making resources scarce. In partially forested areas, though, there would be neither refuge nor resource scarcity, reducing the risk of conflict. To be complete, this project will first test the traditional hypothesis,

Hypothesis 14. Increased forest density is associated with higher risk of conflict.

But, based on the literature and theoretical work done, it is of greater interest whether a more complicated relationship exists. For instance, where forests are completely absent – in desert – may be difficult for both insurgents and government armies. Moreover, from the view that larger populations lead to greater conflict, deserts are also much less likely to be populated with possible conflict actors. On the other extreme, in extremely densely forested areas, the same challenges of survivability may prevent conflict for both insurgents and government actors. Meanwhile, the dense forests are unlikely to contain attractive targets for rebel capture. Thus, for a variety of possible explanations, this project also tests,

Hypothesis 15. There is a parabolic relationship between forests and conflict, with both densely forested and sparsely forested lands associated with a higher risk of conflict.

The outcome of the testing of this hypothesis may not have a clear operational explanation, given the range of options outlined, but it will provide a guide for potential distinctions between types of conflict and it could give shape of future research on the role of forests in conflict.

To make this assessment, though, careful consideration is given to the measurement of forest cover. Rather than simply highlighting the most densely forested areas, the forestry variable in this dataset measures a gradient of the "roughness" from forestation meaning that even low-level foliage is given a score. This construction of the variable will hopefully avoid the same type of problems experienced in some mountainous territory designations where only the most mountainous territory is considered. (The construction of this variable is detailed in the data sources section.)

Natural Resources

It is often presented that, regardless of measurement, natural resources tend to increase the risk of conflict even though one would expect that the discovery of new resources would increase the prosperity of a country – a phenomenon popularly known the "resource curse." This topic has been the subject of a considerable amount of research and focus, particularly since the early results of Collier and Hoeffler's study of resources and conflict began emerging in the late 1990s (Collier and Hoeffler 1998; Reuveny and Maxwell 2001; Ross 2002; Humphreys 2005; Lujala et al. 2005; Lujala et al. 2007; Kahl 2006; Aspinall 2007; Lujala 2010). One of the interesting aspects, which Michael Ross notes, is that the academic study of the role of natural resources in civil conflicts has lagged behind the progress of the policy-making world (Ross 2002, 2004b, 2006). Policy-makers have begun initiatives like the Kimberly Process to control diamonds believed to fuel conflicts and the Extractive Industries Transparency Initiative (EITI) to create transparent accounts for resource revenues, even while academics still do not have a consensus on the precise role that minerals or other extractive resources play in the conflict. Part of the cause for the lack of consensus may in part be a result of the varying impact that natural resources can have on a potential conflict zone. While most discussions emphasize the risk that resources can pose, it is also clear that they have been used as a crucial form of state financing, potentially enabling governments to beat back insurgencies more effectively (or to merely prolong the conflict by financing both sides). For example, although diamonds are often popularly associated with civil conflict, such as by advocacy groups like Global Witness (Witness 2006), and academic findings have been mixed – from studies that find secondary diamonds are associated with conflict onset while primary diamonds are not (Lujala et al. 2005) to findings that suggest that diamonds are generally not associated with conflict onset (Buhaug and Rød 2006). These mixed findings have created an environment in which the academic literature is not entirely clear on the role that diamonds play in conflict. Moreover, there has been some difficulty disentangling the issue of whether resources are the spark that causes civil war or whether they are simply the fuel that allows a fire caused by something else to grow and persist.

One case study illustrates this situation nicely: In December 1983, a violent separatist struggle emerged in Casamance in the southeastern corner of Senegal. Casamance was replete with natural resources, particularly cashew nuts and contraband cannabis – neither of which are included in the disaggregated variables in this study (Humphreys and Mohamed 2005). As a country, Senegal has also been coded by Collier and Hoeffler as being significantly dependent on natural resources at this time, and the coding used by Collier and Hoeffler includes some agriculture resources as primary commodity exports, which are interpreted as “natural resources” (Sambanis 2003; Collier and Hoeffler 2004). On the face of it, this case would seem like a solid instance of resource wealth leading to conflict. However, an examination of the circumstances of the conflict as case study reveals that the initial fighting – the conflict onset – was financed by subscriptions, while the subsequent duration of the conflict was financed by selling of cashew nuts, cannabis and petty crime. (It also does not appear that the Camançais leaders used rhetoric about the potential wealth of the region if it were separate from Senegal; rather, the leaders drew analogies to colonialism and tapped a belief that Dakar had made insufficient investment in the region.) The value of resources was not one-sided, though. The Senegalese army also relied on the selling of timber to finance its operations, allowing it to more aggressively combat the insurgency (Humphreys and Mohamed 2005). This case study highlights a number of challenges in the empirical study of natural resources and conflict: a) differentiating between natural resources as a cause versus an extender of conflict; b) separating the effect of rebel financing through resources from state financing through resources; and c) understanding which resources are most likely to develop conflicts and which resources are most likely to inhibit conflict.

Natural resources, like rough terrain, have been measured in a range of different methods and data specifications, but it often focuses on “primary commodity exports.” This measure is used in a number of key studies (Collier and Hoeffler 2004), but its meaning is often debated, both at an empirical and theoretical level. The variable termed “primary commodity exports” can be composed in a number of different ways. As Fearon demonstrates, some of the strongest proponents of the variable – Collier and Hoeffler – use a version of the variable in their 2004 "Greed and Grievance" article that does not include diamonds but does include coffee, wheat, and oil exports (Fearon 2005). Collier and Hoeffler argue that the opportunity to loot natural resources can start or sustain civil conflicts, a theory sometimes referred to as "resource predation." But, the variable used by Collier and Hoeffler includes a number of resources, such as agricultural products, that are very difficult to loot unless rebels can control state infrastructure or the export trade, as Sambanis notes (Sambanis 2003).

Meanwhile, Fearon and Laitin focus their study on "fuel exports" to address the problem of cash crops interfering with the "primary commodity exports" variable. On the other hand, this approach leaves out some of the most lootable resources, such as alluvial diamonds (Fearon and Laitin 2003b). Perhaps the most damaging criticism of the measure as a whole when used in most aggregate studies is that it generally is formulated as primary commodity exports as a ratio of GDP. This formulation leaves the variable vulnerable to becoming endogenous: as a country approaches, endures, or recovers from conflict, its GDP is likely to decrease, thus making the variable increase as a result (Lujala 2005). Untangling this endogenous effect is not easy – although lagging the variable by a year or more can help in country-year studies, it still does not completely remove endogenous effects, especially in models where lagged negative GDP growth is found to associate with conflict.

There has also been criticism that the "primary commodity exports" variable focuses on the wrong way of measuring the impact of natural resources. Lujala has argued that production, rather than exports, would be a more effective measure (Lujala 2005), and Snyder and Bhavani have noted that whether a resource is lootable, non-lootable, extracted by taxable industry, or non-taxable artisans has more explanatory power than the exports themselves (Snyder and Bhavnani 2005). Naturally, questions of how variables should be addressed tend to be viewed through the lens of the type of research considered: aggregate researchers such as Collier and Fearon tend to emphasize different country-wide measures (particularly as a ratio of GDP) and argue about what should be included as a resource while disaggregate researchers such as Lujala tend to focus on where resources are located and how they could be accessed. Both approaches have merit, but it is important to differentiate between the theoretical explanations that they present. In almost all cases, the country-wide, aggregate assessments of resources tend to provide information on the nature of the state – whether it has a solid tax base, how dependent it is on one or two commodities whose prices may fluctuate, and how valuable the state is as a "prize." On the other side, geographically varying, disaggregated assessments of resources tend to provide more information on how rebels use resources and how "lootability" affects conflicts.

For the purpose of the discussion on natural resources, this section is limited to oil, diamonds and metals, which are discussed separately. While there are other resources – timber, narcotics, cocoa, and others – that have been linked with conflict, they are considerably more difficult to track and they are associated with far fewer conflicts. Thus, while this thesis considers them in a qualitative manner and with occasional references to case studies, there will not be an effort to systematically assess them, given the absence of good data for the whole of Africa over the time period in question on these areas.

Oil

Oil – “black gold” – is often cited as a cause of both internal and international conflict.

Using the country-year approach also leads to results that address only what are perceived as statewide causes of conflict. The most relevant example of this situation for this thesis is the broad academic consensus that oil revenues increase the risk of civil war. This result consistently returns in country-year studies of civil conflict, while similar methodological attempts to demonstrate the role of non-fuel mineral resources in increasing conflict risk have not been as successful (Ross 2002, 2004b; Collier and Hoeffler 2004; Fearon 2005). As Philippe Le Billon and Michael Ross point out, conflicts that seek to take control of a country’s oil resources require taking control of the central infrastructure in order to extract the desired income (Le Billon 2001b, 2005; Ross 2006). As a result, conflicts over oil tend to be statewide or nearly-statewide, meaning that the geography of the conflict more closely matches the geography of the country. While there are cases, such as in Nigeria, Russia, Indonesia (Kaldor et al. 2007), where oil revenues sparked a more localized conflict, these cases tend to be in non-contiguous or large states where there is much more substantial variation in geographic factors (Le Billon 2001b). (There are definitely examples of oil-driven conflicts in smaller states, such as in Nagorno-Karabakh, and the preceding sentence is only a broad generalization.) Moreover, even though the result that oil causes conflict is robust, there is room for some disaggregating of the variable to get a clearer picture of the effects. When the effect of oil is measured, it is often assessed using exports as the focus. While this data is often most available through international trade databases, there is an argument for assessing oil based on production instead, since it might more accurately portray how much total oil – both for domestic and external consumption – is available for rebels to capture and is less subject to the vagaries of price fluctuations in the oil market. (Under the theoretical explanations that oil increases conflict risk by increasing the "prize" of the territory or that oil provides rebel financing, this argument is particularly strong.) In her 2005 study, Lujala focuses on production and differentiates onshore and offshore oil production. Her study finds that "onshore oil production increases the risk of conflict onset by 50% while offshore production has no effect" (Lujala 2005, 2010). This finding makes sense given that offshore oil, which can be miles from a coastline, is virtually impossible to "capture" or access for rebel financing. Also, offshore oil production is almost certain to be taxable by the state, which would provide funding for government forces should it be necessary to defeat an insurgency. Thus,

Hypothesis 16. Proximity to onshore oilfields corresponds with increased risk of conflict.

Hypothesis 17. Density of onshore oilfields corresponds with increased risk of conflict.

The question of the role of oil for this project is not limited to the likelihood of onset, though. Questions of duration and intensity are important as well. On the issues of duration, there is compelling evidence at the aggregate and disaggregate level that oil resources significantly increase the duration of a conflict. In a disaggregate study using the location of oil resources, Lujala et al. find that the presence of oil reserves in a region increases the length of a conflict by as much as 4 times (Lujala et al. 2007). While this dataset is not well-structured to test duration, these findings suggest that oil may make conflict more intense. Moreover, in conflict with oil as a significant component, not only do the combatants have a potential source of funding but, from a game theory perspective, the value of the prize is higher in a governmental conflict, which would be most likely to provide value from the capture of the state oil resources. Thus,

Hypothesis 18. Onshore oil deposits are likely to increase the number of battle-related deaths from a conflict area.

Oil is also an area where there may be a relationship between its location and its impact on conflict. Based on the general dynamics of conflict financing with natural resources discussed above, it seems reasonable that locations where onshore oil is located close to the state capital are more likely to make a source of state financing and less likely to spark conflict as a "prize" for insurgents – whether governmental rebels or territorial secessionists – to claim.

Hypothesis 19. Onshore oil deposits are likely to decrease conflict when they are located near to the state capital and increase conflict when they are located further from the state capital.

Diamonds and other Minerals

Non-fuel minerals are one of the areas where the least consensus has been reached in the literature. There are strong arguments that "primary commodities" are strongly related to conflict with supporting evidence from aggregate studies (Collier and Hoeffler 2004). The data for "primary commodities" is heavily associated with oil, generally do not take into account a broad range of non-fuel minerals, and can include re-exported goods not actually produced in the country (Humphreys 2005).

Mineral resources, particularly those that are easily lootable, theoretically tend to spark localized conflicts that do not spread to the entire geography of the state. Instead, they stay constrained in a smaller territory within the state, which increases the disparity between the variables recorded in a country-year study and the actual variables at the location of the conflict. As a result, one strategy that may produce more effective measurements of the role that mineral resources play in conflict would be a geographically-driven analysis that featured spatial statistics.

More compelling data has come from Lujala et al. who have analyzed the distinction between primary (non-lootable) and secondary (lootable) diamond resources using a country-year model. (This study relies on the same dataset – DIADATA, which is georeferenced – but Lujala et al. produced their analysis before reliable conflict polygons were available.) In their findings, there is a strong relationship between secondary diamonds and the incidence of civil war, but a less clear relationship with the onset of civil war (Lujala et al. 2005). In contrast, Ross finds that diamonds increase the onset of civil war using data generated on the value of diamond product per capita for both primary and secondary diamonds (Ross 2006). (Ross' approach, which includes calculating country-specific values for diamond production, is not replicated here – the only data used is the presence of sites regardless of estimated production value.) There is also debate about the extent to which lootable wealth extends conflict. Fearon and Ross find that "contraband" – which includes "precious gems" – leads to longer civil wars as well because the rebels and/or the government can draw revenue during the conflict, reducing incentives for a settlement and increasing funds available for fighting (Fearon 2004; Ross 2006). In contrast, Humpreys finds that diamonds do not extend conflict because military victory can be achieved relatively quickly, sometimes with the help of external actors (Humphreys 2005). Meanwhile, case studies persistently suggest that diamonds extend war (Smillie et al. 2000; Malaquias 2001; Le Billon 2001a).

To find resolution in these conflicting results, this project attempts to proxy the extent to which diamond deposits were likely controlled by the state by including an interaction variable of the secondary distance between the diamond deposit and the state capital as well as the distance between the secondary diamond deposit and the nearest international border. One of the issues that could have led to such divergent results is unknown question of how much diamond wealth is being collected through illegal means. Although Humphreys includes some data that purports to incorporate estimates of illicit diamond trade (Humphreys 2005), it is difficult to gather reasonably reliable data on illicit behavior commerce. One known proxy, though, might be the distance to the state capital. It stands to reason that secondary diamond deposits are more easily controlled by the state, or associated mining companies, when they are located close to the capital than when they are located far away or near an international where smuggling would be easier.

Thus, there are a series of interaction variables that will be tested to tease out different relationships between the state and diamonds. First, the project will test whether diamonds have an overall relationship with conflict.

Hypothesis 20. Primary diamond deposits inhibit both conflict onset and conflict occurrence.

But,

Hypothesis 21. Secondary diamonds will make areas more conflict-prone to both onset and occurrence.

This hypothesis stems from primary or kimberlite diamonds needing many more state resources to extract their value, whereas secondary or alluvial diamonds require minimal state structure to provide value, making them easier sources of profit for rebels. Second, the project examines the relationship between diamond location and conflict.

Hypothesis 22. Diamonds located closer to the state capital inhibit conflict, but that diamonds located closer to a state border will make conflict more likely in those areas.

While conflict is already expected to be lessened around the state capital, the important distinction is that the presence of diamonds will make conflict even less likely. The regression results in this area, may be difficult to determine, though since there are several cases of conflicts in smaller, diamond-rich nations of West Africa, where the distance to the capital and the distance to the border are both very small compared to larger countries throughout the rest of Africa. Nonetheless, it is necessary to use these interaction variables to make distinctions in the types of operational claims that theories make about the impact of diamonds on conflict.

(It is important to note that this project will not directly address the question of conflict duration. This study does not have dependent variables that speak directly to the length of a conflict, but the regression results on the occurrence of conflict should lead to some insights. Also, since assessing when one conflict begins and another conflict ends is not a precise science, the results on conflict occurrence may be more compelling for those who are concerned with coding the true "end" of one conflict and the beginning of another.)

Regrettably there is little information on the role that specific minerals other than diamonds play at a disaggregate level. This study attempts (for what appears from the existing literature to be the first time as of 2009) to use US Geological Survey (USGS) data on mineral deposits, which includes information on about 305,000 mineral deposits (only about 2700 of which are in Africa) to gauge the role that minerals other than diamonds play in the conflict (*Mineral Resources Data System* 2005). (Ross makes use of the USGS data to augment Humphrey's data on diamonds, but does not use the USGS data on minerals other than diamonds (Humphreys 2005; Ross 2006).) Since there is little information other than a range of case studies that demonstrate the relationships between specific minerals and specific conflicts (Jackson 2003; Sambanis 2003; Ross 2004a), it is difficult to hypothesize a general relationship for these minerals. In general, though, using some of the framework suggested by le Billon, it is possible to sort minerals by their value to rebels groups with minimal mining infrastructure, who would seek valuable and light minerals located nearer the surface, or their ease of control for the government, who would easily control minerals located deeper in the earth or requiring more processing. Thus,

Hypothesis 23. Gold, copper, chromium, platinum and coltan correspond with higher risk of onset and occurrence of conflict, while iron is negatively associated with the onset and occurrence of conflict.

In assessing the results from these regressions on specific minerals, it will be important to take into account the extent to which the minerals are concentrated in only a small number of countries, particularly South Africa and Zimbabwe.

Another aspect of natural resources' relationship with civil war is the effect that it can have on the severity of the fighting. In a largely qualitative and case study approach to conflict, Humphreys and Weinstein reports that the ties to natural resources can make conflicts much more severe, particularly in its effect on civilians (Humphreys and Weinstein 2006; Weinstein 2007). Also, conflicts around minerals can bring in private interests and other countries, as is often cited in the case of the many rounds of conflict in the DRC, which can increase the number of actors wielding violence to achieve their aims in a lawless environment (Global Witness 2005, 2006). Thus,

Hypothesis 24. The presence of mineral deposits, regardless of their type, is increases the number of battle-related deaths.

Population

Population is one of the other areas where there has been significant evidence of a relationship with civil war. Although it is complicated by questions of state size, it is clear that more populous countries tend to have more civil wars than less populous countries (Fearon and Laitin 2003b; Collier and Hoeffler 2004; Hegre and Sambanis 2006). But explanations for the cause relationship of large populations and civil conflict are numerous, and sometimes contradictory. For the discussion of this issue, it is worth separating the statewide population variables – the size of a state's population, density, growth and distribution – from the local population density size, density and growth.

State Population Size, Density, Growth and Distribution

Absolute state population size has often been associated with civil conflict. The theories for this association can either focus on the state or the rebels. From the state perspective, when the population reaches very high levels, the state is not able to address the concerns from all sectors of the population. Moreover, the state will find it more difficult to "penetrate" its own population to put down insurgencies, as it attempts to handle all of the other demands of maintaining a large population (Fearon and Laitin 2003b; Collier and Hoeffler 2004; Sambanis 2003). On the other hand, a larger population might correspond with a larger army with which to resist rebel movements; a theory referred to as "economies of scale in defense" (Hegre and Raleigh 2005). From the rebel perspective, the absolute population of the state might make it easier to find unanswered grievances within a population and therefore to cobble together a coalition of anti-government forces. It is also often supposed that a large state population is likely to be related to heterogeneity (addressed in greater detail later) which could increase the chances of finding social sub-groups that could bond together against the government (Hegre and Raleigh 2005).

Controlling for "social fractionalization," Collier and Hoeffler confirm the positive relationship between state population and conflict (Collier and Hoeffler 2004). In a sensitivity analysis of a number of other studies, absolute population size is demonstrated to have a robust and positive relationship with civil war (Hegre and Sambanis 2006). In Fearon and Laitin's aggregate country-year study, though, they control for the impacts of ethnic and religious diversity, and the results still indicate a positive relationship between state populations and conflict (Fearon and Laitin 2003b). Moreover, Fearon and Laitin note that the log population of a state and its diversity has little correlation, even controlling for

China, which they measure as largely homogenous (Fearon and Laitin 2003b). Thus, the central impact of absolute state population size as measured by most aggregate studies appears to be its hampering the state's ability to react to insurgencies, rather than providing an economy of scale for its defense capacity.

In another instance of testing geographical theories with aggregated information, Collier and Hoeffler calculated Gini coefficients for the population distribution in various countries, resulting in a 0 when the population was perfectly evenly distributed and a 1 when the population was entirely concentrated within a 20km x 20km square (Collier and Hoeffler 2004). The tests by Collier and Hoeffler resulted in consistently negative coefficients, indicating that population concentration reduced the risk of conflict. However, as Hegre and Raleigh note, this aggregate measure fails to capture key distinctions – such as the difference "between a country where the population is concentrated in one cluster covering 10% of the territory and one where the population is concentrated in two clusters of 5% each, but with a considerable geographical distance between them" (Hegre and Raleigh 2005). Thus, while these aggregate measures of geographical factors are interesting as a basis for exploring some theories, they need to be supplemented with disaggregate study to ensure that the results are accurate and robust.

There has been an effort by Hegre and Raleigh to incorporate the state population data into a disaggregated study of conflicts. Regressing data on both the population in a grid cell and the population of the cell's country (along with several other variables) against specific conflict events (rather than conflict onsets), Hegre and Raleigh find that there is no relationship between state population and the location of conflicts (Hegre and Raleigh 2005). Instead, they find that population density – both of the cell itself and the neighborhood – is more significant. This finding, which is determined through combining large amounts of local data with some nation-wide variables and performing a Cox regression, does provide convincing evidence that the broadly robust results that a large population relates to civil war onset (Raleigh and Hegre 2009; Hegre and Sambanis 2006).

A related issue to absolute state population is state population density. In theory, along a similar line of argumentation to absolute population, denser populations are more difficult environments for the state to penetrate and therefore root out insurgencies, particular in cases of urban warfare. Dense populations may also make it easier for rebels to communicate and coordinate with each other. On the other hand, sparely populated areas could make it more difficult for the state to be present throughout the territory, allowing rebels to have safe havens. There are a number of variables that are often associated with dense populations, such as higher GDP, greater infrastructure, and urbanization, and with sparse populations, such as lower GDP, weaker road networks, and rough terrain. If these variables are controlled for in the regression of the aggregate data, though, it should be theoretically possible to discern the impact of population density as separate from some of the other related variables. State population density is not included in the Fearon and Laitin study of civil wars (Fearon and Laitin 2003b), but it was examined in Collier and Hoeffler's study, where they find it to insignificant (Collier and Hoeffler 2004). Collier and Hoeffler did note, though, that prior to controlling for other variables, the states that experienced civil war were nearly three times less dense, on average, than those that did not have civil wars (62 inhabitants per square kilometer in countries with civil war compared with 150 inhabitants per square kilometer in countries without civil war). Based on these earlier findings,

Hypothesis 25. Higher state populations are associated with increased likelihood of local conflict.

And, furthering the idea that state populations are associated with making it more difficult for the state to penetrate into its own territory and populace,

Hypothesis 26. Higher state populations make conflict more likely in the areas far from state capitals.

The reason for testing these interaction variables is to try to provide better context and insight into the operational reasons behind the regression outcomes.

Local Population Size and Density

For practical purposes, differentiating between local population size and density becomes of marginal relevance since most studies that address local population use a grid cell format that fixes the size making measures of "size" equivalent to measures of density (Hegre and Raleigh 2005; Buhaug and Rød 2006). With some cases (Murshed and Gates 2003; Weidmann 2008a; Urdal 2008), it is possible to use local political boundaries instead of grid cells, thereby creating a meaningful distinction between size and density, but these studies rarely reach beyond more than a few countries. Their findings, though, suggest that local population changes are relevant factors in assessing the likelihood of conflict in a given area.

When the local population is large, there is a greater chance that the population will include elements that naturally seek rebellion. Articulated as the "per capita risk of conflict" by Hegre and Raleigh, it would suggest that conflict is more likely when the local population is large, regardless of the state population (Hegre and Raleigh 2005). One might also argue that local population density would allow rebels to communicate and coordinate more easily, along the same lines of argument presented for state population density. Moreover, since more densely populated areas also often represent politically, culturally, or economically vital areas of a country, rebels might focus on controlling those areas in any conflict with a government. On the other hand, if a rebel effort becomes centralized it may struggle to elude government forces, resulting in quick victory or defeat. While this theoretical argument is not an argument against conflict being present in an area of dense population, it suggests that conflicts in dense locations are likely to be short-lived.

A number of disaggregated studies have addressed whether local population density is related to civil conflict, and have generally found that more densely populated local areas are related to at least some forms of conflict (Hegre and Raleigh 2005; Buhaug and Rød 2006; Urdal 2008). In the study by Hegre and Raleigh, the findings were that conflict was generally located in areas of denser population when no neighboring cells were considered.

But when both the density of the grid cell and the density of the surrounding cells were considered together, it appeared that conflict was located in the less densely populated areas that were adjacent to heavily populated areas. In practice, this finding would partially confirm the finding that conflict tended to be located on the outskirts of cities. In the study by Buhaug and Rød of African conflicts, they found population had two contrasting effects on the likelihood of conflict: "territorial" or secessionist conflicts tended to be located in less populated areas while "governmental" or revolutionary conflict tended to be located in more populated areas (Buhaug and Rød 2006). Also, in some recent findings from disaggregate studies based on local political boundaries, certain ethnic configurations and population shifts have been linked with low-intensity conflict. In results that are limited to India, Urdal found that population shifts – particularly changes in the religious composition and the creation of a male-dominated youth bulge – increased the likelihood of conflict (Urdal 2008). Unfortunately, the level of detail of data available for India is unusual, and a similar dataset could not be reasonably created in Africa, particularly over the time period that this project studies. Still, there are some more crude metrics that are available. Using the total population and population density snapshots every ten years from 1960-2000, it is possible to construct local population growth rates. While this data does not provide information on the ethnic or religious configuration, it does give a sense of whether the population increases could be pressing on the limited natural resources of the area. Thus, based on earlier findings, *Hypothesis 27. Areas that experience significant population increases have a higher risk of civil conflict.*

The results indicate that the population had different signs for the two types of conflict identified suggests that a single multiple regression may not be the best approach to measuring the impact of local population density on civil conflict since it can have different signs in different cases. The technique of geographically weighted regression (GWR), described earlier, has the capacity to assign different coefficients to predictor variables at different locations, and it might provide a more accurate understanding of the relationships.

Military Capability

There is little debate in the literature that the size of a state's military decreases the chance for conflict. Although one might argue that a large military is more likely to lead to the creation of disenfranchised groups within the military who would be specialists in violence and could more easily attempt coups and other violent actions against the state, this argument does not play very much in the literature. Hegre and Sambanis find that the absolute size of the military, measured in manpower, has a significant and strong negative impact on the risk of both civil war onset and internal armed conflict (Hegre and Sambanis 2006). This result is independent of considerations regarding the size of the population, and therefore it does not support the "economies of scale in defense" theory directly but rather simply indicates that a large military – regardless of the size of the population – inhibits conflict.

Disaggregated studies discussed the strength of a state's military as a useful measure for its ability to penetrate into its territory, but choose to model that strength with income rather than military expenditures or personnel (Cunningham et al. 2005; Buhaug 2007). The difficulty in using income as a proxy for state strength is that, as has been noted earlier, income is an imperfect measure that leaves open a wide range of interpretations as to its operational significance. For instance, does a high national income mean that a country has a strong military that penetrates rough terrain easily or that it has a well-developed infrastructure of roads and airports allowing a smaller military to govern more widely? Or both? Other studies draw on the concept that military resources are an important factor in where conflict plays out, but focus on rebel resources rather than government resources (Weidmann et al. 2006; Hegre et al. 2007). While it is certainly true that rebel resources play an important role in determining the onset, occurrence and intensity of conflict in given locations, the scoring of rebel strength seems open to possible error as after-the-fact assessments are biased by the outcomes of rebellions. Also, the more relevant (and somewhat more exogenous) question is what causes rebels to have significant resources when compared to the state. The usual suspects for this analysis – a strong ethnic base, rough terrain providing safe havens, proximity to an international border providing havens and opportunities for smuggling income – are assessed in this study. Based on the concept of an exponentially decaying loss-of-strength gradient for state power (Boulding 1962), it is true that increased state military spending may be related to the onset or expected onset of civil conflict, but that should not affect the location of the conflict, presuming that the loss of strength gradient still applies. Thus, *Hypothesis 28. States with high military expenditures are less likely to experience violent civil conflict anywhere in their territory.*

Hypothesis 29. States with high military personnel per capita are less likely to experience violent civil conflict anywhere in their territory.

While this assertion seems like a corollary of the common result that high GDP per capita inhibits conflict, it is a more specific method of showing one of the reasons for the impact of high GDP.

Social Fractionalization

Ethnic Fractionalization

Ethnic fractionalization has been considered as a cause for conflict for centuries. The explanations for how differences between ethnic groups in a society could cause conflict have varied. One explanation is a simple "ancient hatreds" argument that suggests that certain ethnic groups have been and will remain in conflict because of a series of real or imagined injustices between the two groups. Another explanation is that ethnic groupings are constructed based on other social factors, such as wealth or geographic location, and then ethnic distinctions are activated for violence by leaders seeking to gather groups for a conflict. The findings on ethnic fractionalization have been mixed, but have generally indicated that only specific circumstances of ethnic configurations lead to violence

(Horowitz 2000). Collier and Hoeffler find that only situations of "ethnic dominance" lead to violence (Collier and Hoeffler 2004). Collier and Hoeffler explain these situations as cases where the ethnic majority may seek to take advantage of their dominance to extract resources or rents from the minority. Fearon and Laitin are more dismissive of the role of ethnicity in conflicts, arguing that the state strength is the far more significant factor in determining where conflict occurs (Fearon and Laitin 2003b).

More recent disaggregate studies using the Georgraphically Referenced Ethnic Groups (GREG) dataset have led to some intriguing results. In a recent paper, Buhaug, Cederman and Rød examined a series of 35,000 dyad-years that paired an ethnic group with its state capital. Using the GREG dataset, they determined that the demographic size compared to the political center, distance from the capital, living in rough terrain, and political exclusion of ethnic groups made them especially prone to rebellion (Buhaug et al. 2009). (It is interesting to note that while Buhaug, Cederman and Rød hypothesize that the distance from the capital and the roughness of terrain for ethnic groups encourage conflict, they do not test any cost distance or cost path variables that might combine the two issues.) The findings that the increased demographic size of an outlying and excluded ethnic group increases the risk of civil war are generally consistent with findings that two roughly equal-sized ethnic groups present the most conflict-prone configuration. Given that the governing group will likely have some significant size, the very presence of a comparable ethnic group creates risk, which as has been noted in aggregate studies (Collier and Hoeffler 2004; Toft 2007), is only further exacerbated by the excluded ethnic group's distance from the capital.

Assessing ethnic conflict using quantitative data poses challenges, many of which were considered as part of the data specification, but it is worth noting an additional aspect that emerges when forming hypotheses: most ethnic conflict analyses have been qualitative and, as a result, have been able to selectively exclude cases that did not meet the general framework that was established. For example, in Horowitz's lengthy analysis of various components of ethnic conflict, he notes that cases such as Tanzania are excluded because "most groups are small and scattered" while Nigeria is included because its large ethnic groups – the Hausa, Ibo, and Yoruba – all interact at the center of state politics (Horowitz 2000). Because data for this study is drawn from an external source – GREG – and not developed specifically for this study, control over what cases to include and exclude has been essentially forfeited. As a result, hypotheses are developed on much more general terms that might normally be present in a qualitative study. Also, because the amount of detail that can be reliably compared between cases is limited, particularly as the ethnic groups map is static, this study avoids assessing more complicated theories of ethnic conflict, such as certain modernization theories.

The data from the GREG digitization of *Atlas Naradov Mira* provides a set of ethnic group configurations and borders from which variables were generated. In general, it seems likely that – if ethnicity plays any role – conflicts will emerge along points of ethnic friction or tension. The reason for conducting the testing in the following manner is that all conflicts are included and distinguished solely on the basis of their aims and the intensity of the conflict, not on whether the conflict is deemed to have “ethnic” roots or not. Thus, this study will test the following hypothesis:

Hypothesis 30. Conflict onset and occurrence will be more likely in areas with high density of ethnic borders.

And, closely related:

Hypothesis 31. Conflict onset and occurrence will be more likely near ethnic borders.

The more complicated question comes surrounding what ethnic configurations within the territory are most likely to lead to conflict. In some key prior studies, configurations of conflict where there has been a large minority or a small minority have been highlighted as the most likely to create conditions for conflict (Collier et al. 2001). The GREG map data does not provide overall percentages for the populations of the ethnic groups, but it does allow for analysis of the amount of territory occupied by various ethnic group configurations.

The GREG dataset, as rasterized for this study, would also allow for testing a range of interaction variables regarding the territorial size of an ethnic group configuration and its proximity to either the state capital or international borders. While there are a range of possible hypotheses, the most pertinent one – that conflict is likely when ethnic groups stretch across international borders – has already been put forward in the analysis of international borders. Other analyses are likely to be more situation specific, possibly with a role for the form of government, and would require a significant number of hypotheses, interaction variables, and additional data specification to test properly. As a result, it will be left to future research.

Country Size

In theory, the larger a country is the more difficult the territory is for a government to administer. On the other hand, the smaller that a county is the easier that it may be for rebels to take control of the whole state. The variable of country size as measured by territory is not included in the major aggregate Fearon and Laitin or Collier and Hoeffler studies (Fearon and Laitin 2003b; Collier and Hoeffler 2004). Since most disaggregate studies have not included country size as a variable, this section will focus mostly on the findings of disaggregate studies, which have tended to focus on three types of distance – distance from the state capital, distance from an international border, and cases of non-contiguity. Non-contiguity is only briefly discussed because it will not be controlled for in this project's dataset.

Proximity to the Capital

Every government faces limited resources and projecting its power and authority across large areas becomes more difficult as the area grows. These state resources are likely to be most concentrated in the capital and then their concentration lessens further from the capital. As state resources become more scarce both “greed” and “grievance” scenarios are possible: with weak state presence, there are more opportunities for rebels to take advantage of the situation, and with weak state presence, there is likely to be relative underinvestment in development in the region. Moreover, once a conflict begins, insurgent groups will likely benefit from the difficulty of projecting state power into the periphery. In the literature, this concept is often referred to as the “loss of strength” gradient which may exponentially decay a state’s power as it departs from its political center (Boulding 1962; Buhaug 2007). Under both scenarios of “greed” and “grievance,” the argument is that conflict will be more likely further from the capitals of countries. This conclusion is broadly supported by the literature, which consistently finds a positive relationship between distance to the capital and conflict (Buhaug and Rød 2006). Thus,

Hypothesis 32. Territorial conflict is more likely in areas further from the state capital.

While, in contrast, because the capital is the seat of state power, which is the target prize for governmental conflicts and wars,

Hypothesis 33. Governmental conflict onset is more likely in areas closer to the state capital than territorial conflict onsets.

This dynamic of the relationship of the conflict with the center of state power has been a key dynamic that differentiates territorial and governmental conflicts.

Proximity to an International Border

While proximity to an international border and proximity to a capital are often negatively correlated, the theoretical underpinning for the idea that conflict should be greater near international borders is different. An international border marks an area where quick export of natural resources is possible and where safe havens may be found in sympathetic neighboring countries. Also, a number of international borders, particularly in former colonies, cut through ethnic groups that may seek to unite their territory – such as the Kurds in northern Iraq and southeastern Turkey or the Tutsis in eastern DRC and Rwanda (Prunier 2004; Salehyan 2010). Despite these theoretical findings, there is little evidence from disaggregate studies that conflicts tend to be located particularly close to borders (Buhaug and Rød 2006). One reason for the strong case studies and limited empirical evidence is that ethnic, economic, and political data have not been used significantly in these studies. Since the case studies that support the idea that international borders matter center on their relationship to smuggling, common ethnic groups, or safe havens, it makes sense to test three interaction elements.

To test smuggling,

> *Hypothesis 34. Conflict is more likely in border areas where there is a high concentration of secondary diamonds.*

To test safe havens based on ethnic groups, *Hypothesis 35. Conflict is more likely in areas that have ethnic group configurations whose territory crosses an international border.*

To test safe havens based on terrain,

Hypothesis 36. Conflict should be more likely in border areas that have rough terrain.

Granted, it is possible that a positive relationship between border-crossing ethnic groups and conflict indicates that the ethnic groups provide safe haven, but it is an important risk factor if statistically significant. Moreover, since the operational relationship between rough terrain is not clear – there are other viable explanations of the role of rough terrain – this last hypothesis is open to multiple interpretations from the data. Nonetheless, it is the best match of data to a hypothesis possible, and it tracks with several other authors who have suggested that rough terrain provides a safe haven for insurgents (Fearon and Laitin 2003b).

One aspect of the capital-border relationship that has not been tested is an effort to combine the two variables. In other words, while there have been assessments of the distance from a grid cell to the capital or a grid cell to a border, there have been no assessments of the relative distance of a grid cell to a capital and a border. Thus, another variable is included in this thesis – the distance of a grid cell from a capital minus its distance from an international border. So, for small countries, the number will be fairly low since any point in the country is not that far from a border or the capital, but in countries such as Algeria and Morocco, whose capitals are located near international borders but on the far edges of the country, this measure provides a distinction between cases when the grid cell is on the border next to the capital and on the border far from the capital. Since most literature indicates that the capital, as the center of state power, is less likely to see conflict, this project tests the following hypothesis:

> *Hypothesis 37. The capital-border distance difference increases so that locations near capitals and far from borders are least likely to have conflict while those near borders and far from capitals are most likely to have conflict.*

Non-Contiguous States

While they do not address territorial size, Fearon and Laitin include a dummy variable for non-contiguous states under the theory that non-contiguous states have natural boundaries that separate the locus of state power from at least one area of its territory. The concept of country size appeals to the same principle that state power will dissipate as it gets further from the capital. They find that the non-contiguous states have a higher risk of civil war onset but the findings were not significant to a 95% confidence level (Fearon and Laitin 2003b). In this study, the presence of a non-contiguous state would be most important for the section of territory that does not include the state capital in its borders. In Africa, there are a few relevant examples of this situation: Zanzibar in Ethiopia, Cabinda in Angola, the Continental Region of Equatorial Guinea, and the various island nations, such as the Comoros. These cases, though, are quite limited, and it does not appear that there are enough cases to justify establishing a dummy variable for those grid cells beyond the existing variable for distance to the state capital. A dummy variable would just represent this limited number of cases, and could appear unduly significant or insignificant based on only a few instances. If this project were expended beyond Africa in a future research project, then more robust and significant findings from non-contiguity might be possible.

Form of Government

Democracy, Anocracy, Autocracy

Democratic governance – defined as a broad range of civil and political liberties – has some elements that would suggest that it is better at channeling opposition movements towards the political system and away from violence. For example, in democracies with open elections – or even partly open elections – formerly armed groups such as FMLN in El Salvador or Partai Aceh in Indonesia have transformed themselves into political movements to gain power through the ballot box. For these groups, the opportunity to gain power through political action often requires some transformations in their structure but it can be less costly than facing a government-supported army. Conversely, as Collier and Hoeffler describe, autocratic governance increases the risk of backlash in the form of civil war. This theory is supported by the data that they collect and analyze from the Polity IV dataset (Collier and Hoeffler 2004). There is also data that suggests that governments that are stable – be they autocratic or democratic – are less conducive to civil war than those that are inconsistent, middling, or “anocratic.” Fearon and Laitin fail to find any significant correlation between democracy and civil war, but they do find a relationship between "anocracies" and civil war (Fearon and Laitin 2003b). Collier and Hoeffler investigate including the autocratic score as well as a measure for democratic governance, but find no relationship for the autocratic score. Fearon and Laitin also try a number of different scorings for governments, but find that the Polity IV scoring system is the only one that reaches significance. It is worth noting that Fearon and Laitin found anocracies reached significance only when regime transitions were coded as anocracies – when the values for regimes were instead interpolated, the significance of anocracy went down slightly, failing to meet a 95% significance threshold, while the significance of political instability increased in their model (Fearon and Laitin 2003b). This general conclusion that states with middling regime types are least stable and most prone to conflict has been noted by a range of scholars (Gurr 1974a; Huntington 1991; Ellingsen and Gleditsch 1997; Hegre et al. 2001b; Sambanis 2001; Bremmer 2006). Based on the existing literature,

Hypothesis 38. States with anocratic regimes will be more likely to experience conflict onset and occurrence in their territory.

And, correspondingly,

Hypothesis 39. Autocratic and democratic states will be less likely to experience conflict onset and occurrence in their territory.

One of the complications suggested is that the relationship of the political system to conflict may be partially endogenous because violence can be involved in scoring the political system. In democratic regimes, violence can scored against the openness of the regime. As a result, mobilized opposition in autocratic regimes boosts its "democracy" score – meaning that precursor violence may be drawing both autocracies and democracies towards scoring in the middle as an anocracy (Hegre et al. 2001b).

Based on these scoring complications, it is difficult to draw a definitive conclusion, but the key element appears to be the measurement of the political system and the scoring of the conflicts. As a result, the straight Polity IV score and a transformation of the score to emphasize anocracies will be used.

Recent studies have worked to refine the finding that democracy produces a more stable and less conflict-prone polities by disaggregating democracies into multiple categories – some categories have included majoritarian, proportional, candidate-driven or party-driven (Reynal-Querol 2002; Carey 2007). While these efforts have disaggregated the regime type, there have not been as many efforts to disaggregate the territory of the state. One area that has not been significantly explored is the relationship between different locations within a country and its political system. In other words, because most of the past studies have focused on the nature of the political regime at the center and whether the country as a whole experienced conflict, there has not been much study of how political regime affects the propensity for conflict near the capital or in the periphery. (One notable exception is some work on the modeling of loss-of-strength gradients by Halvard Buhaug (Buhaug 2007).) And, because geographic studies have focused on such wide ranges in time, political data has generally not been included.

One would ideally want to have data on both the central and local political regime in assessing the impact of political structure on conflict. It seems likely that the risk of conflict from political regime is related not only to the political type of the central government but also the local government. For example, are provincial or state governors elected locally or appointed by the state executive? Are there local legislatures that are elected? Unfortunately, there is not a reasonably reliable and comprehensive database of local regime types covering the period of study for this project – but that does not mean that an initial effort using only data on central regime type cannot be pursued.

Based on the existing information and research, though, it seems that a few tentative hypotheses could be drawn about how the central government regime might impact rebellion. For instance, from past research that governmental conflicts are most likely near the capital of a state (Buhaug and Rød 2006), and it seems most likely that governmental conflicts or short but violent coups might be most common under autocratic regimes where the entire state can be captured simply through changing the central leadership. Thus,

> *Hypothesis 40.* *The proximity to the capital and the political regime type have an interaction effect that makes conflict more likely near the capital in autocratic regimes.*

One could argue the contrary – that autocratic regimes, which are often characterized by concerns about instability, would be most likely to lock down the capital to prevent coups or changes in government, forcing conflicts to be waged further from the capital. Based on previous findings, though, it seems less likely that governmental conflict operates in that fashion.

Conversely, in democratic regimes, it is both more likely that democratic institutions exist at the local level and that those democratic institutions allow local power bases to develop away from the capital. As a result,

Hypothesis 41. Territorial and governmental conflict are more likely to occur far from the state capital in democratic regimes.

The opposite argument could have some validity – that because local democratic institutions are more likely democratic regimes, the local interests are more likely to be placated by capturing local power without needing to fight the state, meaning that only those with no meaningful local power base to capture (i.e. those near the capital) would need to challenge the government. Also, in democratic regimes, it is possible that the forces near the state capital could be of a significantly different ethnic and political make-up than the state overall, putting the local inhabitants and the government members in potential conflict.

It is also sometimes argued that in open, elected democratic governments there is a greater tendency to avoid violent reactions against opposition groups unless completely necessary. The electoral and political costs of acting violently against a state's own population are believed to be higher in democratic regimes. On the other hand, the delay in responding may allow opposition movements to grow larger, and the ability to actively organize political opposition in most democracies could allow for much larger conflicts when they occur.

Conclusion

The literature of the past several decades, and the last decade in particular, has thoroughly pored over the information about possible aggregate explanatory variables for civil conflict, and recent efforts have opened the door to considering disaggregate factors. This thesis will now push to integrate the two. The overall aim of this series of different hypotheses developed from the literature review is to integrate the theories that have emerged in the aggregate and disaggregate analyses into testable propositions using data from both the state and local level. Taken collectively, this set of hypotheses offers insights into the broader question of whether conducting analysis that distinguishes between territorial and governmental conflict, as well as low-intensity conflicts and civil wars, is a worthwhile pursuit for the field.

In the chapter that follows, as the data regressions are performed and reviewed, the aim will be to ensure that the analysis continues to push for integrating aggregate and disaggregate information in the analysis. For example, rather than simply focusing on whether democratic, anocratic, or autocratic regimes are associated with more civil conflict, the analysis pushes to ask what parts of territories of these different regimes are most vulnerable. Similarly, rather than simply focus on whether proximity to a state capital is associated with more civil conflict, the analysis pushes to understand whether high income or low income states will accentuate the importance of the state capital's location in conflict patterns. These analyses suggest the relationships between aggregate and disaggregate factors in a manner that has not been fully explored in past literature.

Finally, these hypotheses take advantage of a wide-ranging dataset to explore the data for both correlates of conflict onset and occurrence. Considering both onset and occurrence provides insight into two important conflict dynamics – the start and continuation of conflict – but it also speaks to the broader aim of understanding whether different types of conflict exhibit significantly different correlates. This set of hypotheses outlines a path for examination not only regarding similarities or differences between types of conflict onsets but also types of conflict occurrence. Thus, when considering whether types of conflict are different in their correlates, suggesting that types of conflicts need to be separated to be effectively studied, the assessment will go beyond conflict onset and include continuation of the conflict as well.

CHAPTER 5: ANALYSIS OF RESULTS

Note: In the follow analysis, the word "hostilities" is used to describe findings or categories that apply to both low-intensity conflict and civil war. The reason for introducing the new term is to help remove some of the confusion between "conflicts" as all hostilities and "conflicts" as "low-intensity conflicts" with 25 to 999 battle-related deaths per annum.

This chapter analyzes the results of extensive regression analysis of the dataset constructed for this project. At the outset of this project, I outlined five broad aims for the research:

1. **Combine Aggregate and Disaggregate Data:** Is there a reasonable method for including data traditionally associated with aggregated country-year studies, such governance or economic data, into disaggregated studies, which have traditionally focused on geographic and environmental variables with no temporal variance?
2. **Differentiate Conflicts by Aims and Intensities:** What analytical insights are achieved by differentiating conflicts by their aims (territorial secession versus government capture) and their intensity (low-intensity conflict and civil war)?
3. **Evaluate Specific Relationships for Variables:** By synthesizing aggregate and disaggregate variables using geo-referenced data, what insights can be determined on the relationships between political, economic, social and geographic variables and conflict emerge? Some of the major considerations include (but are not limited to):

   a. Does a pattern emerge that population density is a contributor to the risk of civil conflict in some areas and a deterrent in others?
   b. Does a pattern emerge that rough terrain (though either mountains or forests) are contributors to the risk of civil conflict in some areas and deterrents in others?
   c. Does a pattern emerge that mineral deposits (both fuels, such as petroleum, and non-fuel, such as diamonds) are contributors to the risk of civil conflict in some areas and deterrents in others?
   d. What role do borders between ethnic groups play in the likelihood that conflict will onset?

4. **Consider Onset and Occurrence:** What role do variables relevant for conflict onset play in the continuation of civil conflict? Do conflict onset and conflict occurrence have similar or distinct causes?
5. **Determine Validity of Distinctions Between Conflict Types:** Is there a distinction between the correlates of different types of conflict for the onset and occurrence of hostilities?

The first of these questions has mostly been addressed in earlier discussions of the methodology that outlined a hybrid approach to bringing together aggregate and disaggregate data in grid cells tracked by five-year periods for a three-decade span. Although there have been conceptual discussions of the final four sections of the research agenda in earlier chapters, this chapter will address these four sections with the results of the analysis conducted. The chapter discusses the results of several regressions that compare different types of conflict, and most comments refer to the findings of two "base" regressions that are summarized in a table at the end of the chapter. I conducted several additional regressions that reflect the robustness of the findings, and those tables are also included at the end of the chapter.

The broadest and most fundamental question posed in this thesis is the final one: Is there a distinction between the correlates of different types of conflict for the onset and occurrence of hostilities? From the research and analysis conducted in this project, it is clear that conflicts with different aims – whether territorial secession or government capture – and different intensities – whether low-intensity or war – display distinct correlates. For example, based on the regressions conducted for my thesis, governmental and territorial hostilities demonstrate distinct economic conditions. Governmental hostilities correlate with lower statewide GDP growth, while, in contrast, territorial hostilities are more likely to have higher statewide GDP growth.[17] Governmental and territorial hostilities also exhibit different environmental conditions: governmental hostilities are much less correlated with local environmental factors while territorial hostilities are more likely to see lower levels of local rainfall and lower local forest density.[18] Similarly, on social factors, governmental wars have little correlation with ethnic border density while territorial hostilities correlate with ethnic border densities, although the sign (+/-) depends on the intensity of the conflict. In demographics, higher total state populations were consistently associated with territorial hostilities at all levels while total state population had negative significance only in the case of low-intensity governmental conflict. Similarly, on military issues, increased military expenditures consistently had a negative correlation with

[17] Finding based on regression using local rainfall.
[18] Finding based on regressions for local rainfall and local forest cover conducted separately (see tables at end of chapter).

territorial hostilities at all intensity levels, while military expenditures had a positive relationship with low-intensity governmental conflict. Despite numerous differences between the correlates of governmental and territorial hostilities, there are some examples of variables that do have a consistent impact in all cases: local rainfall, local forest cover, proximity to an international border and Polity IV score all hold the same sign in each case where the variable is significant. These findings reinforce the view that governmental and territorial conflict are different phenomena with distinct statewide and local correlates that cannot be completely explained in a regression including all types of conflict as previous studies have done.

Conflicts with different intensity levels also displayed distinctions that suggest that they ought to be approached as separate phenomena. Other than local elevation range, no variable held significance and the same sign in regressions for general civil war and low-intensity conflict. The findings are sufficiently divergent in some areas that it suggests that differentiating conflicts by their aims and intensities may provide more useful and focused results. The results are compelling and explain some of the distinctions between the different types of violent civil conflict that states experience. The analysis also broadens the lens beyond the traditional focus on simply wars – those with more than 1000 battle-related deaths per year – to include a wider range of conflicts with different associated factors. Ultimately, this chapter aims to provide an initial guide to these conflict factors that could be continued with further empirical work or the inclusion of case studies in future research.

While a major consideration of this project is the differentiation of conflicts by aims and intensity, the chapter is structured along onset and occurrence with intensity assessed through the lens of both. The reason for this construction is that onset and occurrence require different methodological approaches to assess, while aims (classification as territorial or governmental) and intensity (classification as low-intensity conflict or war) are included in both approaches. A key aim aim of my thesis is to better understand the distinctions between governmental and territorial conflicts as well as between low-intensity conflicts and wars.

Analysis of Onset

Note: Onset is defined as the first year of a conflict in a given category. In other words, if a conflict starts as a low-intensity conflict but ultimately becomes a war, then the conflict is scored as a war in its first year. There are only two cases in the dataset where this issue arises: Chad between the 1970s and 1980s (for a conflict that onset in 1965) and Burundi between the early and late 1990s. The situation in Chad is particularly complicated because there are multiple conflicts occurring in that territorial space at that time. In all other cases, the conflict remains at its intensity level throughout the periods in which it is active.

To analyze the civil conflict data, a framework approach was used that assessed the attributes associated with different types of conflicts – whether governmental or territorial – and different levels of violence – whether a conflict (25-999 battle-related deaths per year) or a war (1000+ battle-related deaths per year. The result was a series of logistic regressions that provided an overall assessment of what variables were associated with conflict overall and then which variables had particular relevance for sub-types of conflict.

This analytical framework helps to distinguish between the correlates of war for a variety of different types of conflict. It also assesses what variables are associated with less intense conflict and more intense war without engaging in a linear regression to estimate the precise battle-related deaths per conflict. Given the lack of precision in much of the battle-related death data, it seemed more reasonable to simply assess intensity based on the two broad categories of conflict and war rather than at the level of specific numbers of battle-related deaths.

Totals n = 95,886 inc. 8761 **onsets** (9.14% onset)	Governmental Hostilities n = 6404 cells	Territorial Hostilities n = 2357 cells
Civil War n = 6039 cells	Governmental War n = 5090 cells	Territorial War n = 949 cells
Lower-Intensity Conflict n = 2722 cells	Governmental Conflict n = 1314 cells	Territorial Conflict n = 1408 cells

From a basic overview of the conflict data, it emerges that governmental hostilities are much more likely to onset as war rather than lower intensity conflict, while the reverse is true for territorial hostilities, which are much more likely to onset as conflict than higher intensity war.

A summary table of the overall findings for onset of violence is included below. To review the findings, the analysis will divided into different thematic areas where the findings between the regressions are compelling. This approach should be clearer than approaching the findings by the individual types of conflict.

Findings for undifferentiated conflict not consistent with sub-types

When the regression for all kinds of hostilities – governmental and territorial, conflict and war – were included in a logit regression, the results were not consistent with the results for the sub-types of conflict. There were several variables – such as GDP growth, proximity to international borders, or change in local rainfall per capita – that were significant for the overall regression but significant with different signs or insignificant for key sub-types of conflict. Thus, while the overall regression shows both local and statewide variables at play in conflict, these results did not correspond with the results for many of the sub-types of conflict. These results also highlighted that the role of international borders is important, but it did not help to differentiate between the possible explanations for reasons why borders increase onset of armed hostilities. It is surprising that some of the more robust findings from the aggregate civil war studies – such as the significance of the size of the country, size of the population, or GDP per capita – were not found in this overall regression. The most likely explanation for the absence of these overall findings suggests that the state-level variables were not as significant at the local level, which is where this study focuses.

In the next several sections, I discuss some of the findings for the broader types of conflict and what they may indicate about the relevance of certain variables and theories of conflict. Before delving into those issues, though, I want to highlight the broadest and most important point that has emerged from my research: that there are significant differences between different types and intensities of conflict, sufficient to suggest that a single empirical explanation that satisfies all types of conflict is not reasonable. Moreover, efforts to divide conflict into groupings such as "ethnic" conflict or "son of the soil" conflicts to modify a general theory are also not sufficient to explain the differences experienced at the local level. To summarize, using the base regression from this dataset:

Variables that are **both significant and have <u>opposite</u> signs for <u>governmental</u> and <u>territorial</u> conflict**	Variables that are **both significant and have <u>opposite</u> signs for civil <u>war</u> and <u>low-intensity</u> conflict**	Variables that are **both significant and have the <u>same</u> signs for <u>governmental</u> and <u>territorial</u> conflict**	Variables that are **both significant and have the <u>same</u> signs for civil <u>war</u> and <u>low-intensity</u> conflict**
1. Local population density 2. Change in local rainfall between periods 3. Statewide GDP growth	None	1. Polity IV score (thesis-adjusted)	None

Note: excludes lag variables and constants

So, there are three variables that have significant and opposite signs between governmental and territorial hostilities and there are no variables that were significant for both civil war and low-intensity conflict. These findings are in stark contrast to the findings of Fearon and Laitin, who found variables such as the GDP per capita, state population, percent of mountainous terrain, and "new state" to be consistently significant with the same sign (Fearon and Laitin 2003b). Similarly, Collier and Hoeffler find significant and consistent findings for primary commodity exports, male secondary schooling, democracy, and state population (Collier and Hoeffler 2004). The findings in my thesis suggest that these previous statewide studies may have been insufficient to explain local-level conflict, particularly territorial conflict.

Moreover, based on the regressions that were conducted for my thesis, it may well be that there is not a single regression that can explain the many types of conflict using the same set of variables. In the regressions, many variables were significant to one set of conflicts but not another. For example:

Variables that were significant to **governmental hostilities but <u>not</u> territorial hostilities**	Variables that were significant to **territorial hostilities but <u>not</u> governmental hostilities**	Variables that were significant to **low-intensity but <u>not</u> civil war**	Variables that were significant to **civil war but <u>not</u> low-intensity conflict**
1. Local elevation range 2. Local density of petroleum 3. Distance to state borders	1. Local level of mean rainfall 2. Local forest level compared with state forest level 3. Local mineral density 4. Local distance to secondary diamonds 5. Distance to state capital 6. State military	1. Local elevation range 2. Total state population 3. Oil as dominant commodity export (dummy) 4. Polity IV score (thesis-adjusted)	1. Change in local rainfall between periods 2. Local population density 3. Density of local petroleum 4. Distance to state border 5. State military expenditures 6. State GDP growth

	expenditures 7. State military personnel 8. Oil as dominant commodity export (dummy) 9. Density of ethnic borders		

Note: excludes lag variables and constants

The table above reflects the base regression used for this thesis, but depending on the variables included, the exact variables that are included or excluded change slightly. At the end of this chapter, a few variations on these tables are included to demonstrate some of the more severe changes that occurred when small alterations were made to the variables used. For example, one of the variants on the base regression that uses just the Polity IV democracy score (with modifications) has five variables (per capita local rain change, local population density, distance to secondary diamonds, state military expenditures, and state GDP growth) that are significant but with different signs for both governmental and territorial hostilities. What remains robust is that there are significant differences between the types and intensities of the conflicts. In the following section, I will attempt to explain some of these distinctions, but the broader point holds that these conflict types and intensities have distinct correlates.

State-level GDP per capita not significant to local conflict

In the existing civil war literature, it is generally accepted that higher GDP per capita level is negatively associated with civil war onset (Fearon and Laitin 2003b; Collier and Hoeffler 2004; Hegre and Sambanis 2006; Hoeffler 2009). In the regression analysis, the most surprising non-finding is that state-level GDP per capita has very limited significance – the variable was only negatively associated with territorial low-intensity conflict. This finding is not a function of the combining of different types and intensities of hostilities because state-level GDP per capita is not significant in any of the sub-types regressions that include only civil war (e.g., conflicts with more than 1000 battle-related deaths per year), which is the class of conflict most commonly used to regressions that highlight the significance of GPD per capita. What is more likely as an explanation of the non-finding on state-wide GDP per capita is that the study focuses on conflict at a local level, while previous studies that have highlighted the significance of GDP per capita have generally been aggregate studies that did not specify the portion of the state suffering civil war. Thus, while GDP per capita may be a relevant explanatory variable for whether a state will experience civil war in general, it is not as useful at the sub-national level as originally hoped.

Local and state economic conditions in governmental and territorial conflict

One of the starkest distinctions between governmental and territorial hostilities is the distinct economic conditions under which each occurs. Given the limited findings on statewide GDP per capita, it is particularly interesting to see the local economic factors that are associated with hostilities. On the state level, the state GDP growth was negatively correlated with governmental hostilities onset and positively associated with territorial hostilities onset. At the local level, the *change* (or difference) in the level of rainfall per capita between periods is also negatively associated with governmental hostilities onset, while territorial conflicts exhibited a more robust relationship with having lower levels of rainfall and lower levels of local forest cover. (Territorial hostilities also demonstrated a positive correlation with changes in rainfall per capita, but the relationship was less robust than for governmental hostilities.) If changes in local rainfall are proxies for changes in the local economic conditions and overall levels of rainfall or vegetation are proxies for growth potential, then this finding suggests that economic growth is a significant factor in the risk of onset at both the state and local level for both governmental and territorial conflicts – but with opposite signs. In other words, the riskiest locations for governmental hostilities onset are those areas of negative growth states that are also experiencing declining rainfall, while the riskiest locations for territorial conflicts appears to within higher growth states.

In Miguel et al.'s study on economic growth and civil conflict in Africa, there is a positive demonstrated relationship between local vegetation levels and the level of economic growth (Miguel 2004) using data from Africa. Miguel associates vegetation levels with rainfall in his article, and, at least initially, it appears to be a reasonable assumption that areas with higher average rainfall would experience higher vegetation levels, which are associated with higher GDP growth. (It is noteworthy that while Miguel's study is often described as a study on local rainfall (Jensen 2008), it is in fact a study on local vegetation – a close correlate of rainfall, but not indicative of the same variable under consideration.) The findings from my own analysis suggest that while the negative "economic shocks" that Miguel highlights as being linked with civil conflict risk, particularly for poorer, less democratic countries, only hold robustly with governmental conflicts. Territorial conflicts appear to follow a different set of correlates at the local level.

If the local level of rainfall is taken as a proxy for local economic growth in this project, these findings combined suggest that governmental hostilities emerge in local areas that are shrinking economically while the state is also experiencing less economic growth. In contrast, territorial hostilities were more common in growing states with more arid conditions. In short, local areas most likely to experience government hostilities onset are experiencing economic weakness while the state overall is experiencing weakness, suggesting a "grievance" explanation for the onset. In territorial hostilities, though, the local area has higher economic growth potential – reflected in the higher percentage of vegetation in territorial conflict zones – while the state overall is experiencing strength, suggesting a "greed" explanation for the onset.

Parsing these individual findings on variables is not easy in the abstract, so it is helpful to highlight the cases that may be helping to drive the findings. For the governmental onsets, the most extreme cases with state GDP growth and change in rainfall per capita that are more than one standard deviation below the mean are portions of Congo in the early 1990s. Cases with values that are half of a standard deviation below the mean for state GDP growth and change in rainfall per capita include portions of Chad and Uganda in the early 1980s and portions of Angola in the early 1990s. Of the three, the starkest example of the impact of rainfall is in Chad. The Chadian conflict in the 1980s occurred in a country where over 80% of the country survived on subsistence agriculture (*Chad: A Country Study*) and, according to Sambanis, "the regional division and the competition between Northern Saharan tribesmen against Southerners" was a key cleavage, one worsened by agricultural patterns including drought due to steep declines in rainfall

(Sambanis 2004). In all three conflicts, short-term political factors are also key factors, but the changes in rainfall may have laid the groundwork for more conflict-prone dynamics.

One caveat should be delivered about associating mean rainfall levels as a measure of economic growth in the context of this study: average state-wide levels of rainfall do not have a strong direct relationship with state GDP growth. In fact, none of the metrics that were developed for rainfall suggested a particularly strong relationship with state GDP growth in the data that we collected; none of the metrics had more than r = 0.04 with statewide GDP growth, and both mean local rainfall and the statewide mean rainfall had negative correlations with statewide GDP growth. This situation could be explained in several ways: either 1) mean rainfall explains local economic growth but that growth is not fairly measured in the available data on state GDP growth, or 2) the mean rainfall levels in fact indicate another factor, such as the density of forests (where mean local rainfall and local forest density are correlated at over 0.8), which may have negative correlation with conflict risk. Ultimately, with few local estimates for economic growth, it is difficult to directly connect rainfall to local GDP growth. As Collier noted in 1999, "Since the 1960s, there has been a deterioration trend in African rainfall. However, there is too little evidence to infer its effect on growth" (Collier and Gunning 1999). Moreover, without engaging in an empirical analysis, Bloom and Sachs suggest that the variability of rainfall in Africa – regardless of its average – can be a greater depressor for economic growth (Bloom and Sachs 1998), and, Sachs even continues to directly link rainfall as the prime driver of conflict in Darfur (Sachs 2005). In an empirical study, Landau broadly finds no significant relationship between rainfall and economic growth or associated factors (Landau 1986). Landau suggests that the reason that rainfall was not significant is because rainfall patterns have already had an impact on growth through population settlement patterns, so there is no longer a significant impact in spans of a few decades in the late 1900s. So, despite some civil conflict literature associated conflict and rainfall, there does not appear to be enough evidence in the dataset for this project or in the work of others to immediately draw the connection between rainfall and economic growth.

If local rainfall is not taken as a proxy for local economic growth, what explains the relationship found between local rainfall and a number of types conflict? As noted above, one explanation might be that rainfall and forest level are closely associated, and that forests and/or rainfall are significant to governmental hostilities, territorial hostilities and hostilities overall. Thus, the appropriate lens through which to view the impact of local rainfall is not as strict GDP growth but rather as general resource availability. Even if rainfall cannot be tied directly to economic growth, it can be tied to the availability of basic resources, such as drinking water (from rain), improved forestation, and improved food availability. Further, the lack of rainfall can be tied to other problems, such as increased water contamination, under-nutrition, and public health challenges like diarrhea (Clover 2003). In other words, while local rainfall and local forest cover may not connect with statewide economic GDP growth statistics, it is reasonable and well-founded that they often connect with improved quality of life and reduce some of the greatest health challenges. Hence, whether local rainfall or local forest cover are representative of local GDP growth or just local resource availability, the basic concept laid out earlier regarding the differences between governmental hostilities and territorial hostilities with regard to these variables remains salient. Governmental hostilities reflect a “grievance” dynamic in which the statewide economic growth figures are weaker and the local rainfall experiences a significant negative change. Territorial hostilities, in contrast, reflect more of a "greed" dynamic in which statewide economic growth figures are stronger and the local forest cover (possibly representing local resource availability) is higher than the state average, providing a better opportunity for territorial capture.

This broad division is not intended to ignore what might be legitimate evidence of "greed" elements in governmental conflicts or "grievance" elements of territorial conflicts. The lower level of mean local rainfall compared with higher GDP growth in territorial conflicts could well be evidence of communities being left-out of statewide growth, inciting grievances and territorial secession movements. Similarly, the lower GDP growth figures combined with higher local densities of minerals and petroleum in governmental wars could indicate "greed" components in governmental hostilities. The broader point – which is the primary contribution of this thesis project – is that the territorial and government conflicts express significantly different correlates, suggesting that there is not a single unified theory of conflict that can be empirically demonstrated.

Conflicts and wars

In the regression analysis, there is no variable that is significant for both conflicts and wars; in fact, the distinction appears to be asymmetric. The variables associated with conflict all appear to be related to reducing the potential for the conflict to grow, rather than factors that might be linked to inciting the conflict in the first place. Meanwhile, the variables associated with civil war appear to be linked to attributes that would make a conflict notably more severe by reducing the friction for groups to interact and increasing the potential grievances. I must note here that civil wars were more likely to be governmental than territorial, which explains some of the similarities in their findings. For example, low-intensity conflict had significant correlates with:

1) increased elevation ranges, which might actually prevent the kind of open warfare needed to achieve high casualties;

2) lower total state population, which also reduces the number of available combatants and casualties;

3) oil as the dominant export, which increases the state's capacity to fund itself independently to defeat potential insurgents; and

4) higher Polity IV score, which may indicate that democracies are less tolerant of high casualties and prevent drastic situations from occurring.

In contrast, civil wars had significant correlates with:

1) negative changes in rainfall per capita and negative state GDP growth, which could both increase grievances against the state and decrease the opportunity cost of conflict for insurgents;

2) increased local population, which increases the number of available combatants and potential casualties;

3) density of petroleum deposits, which increases the prize for success in the conflict, particularly statewide conflicts; and

4) lower state military expenditures, which indicates a more limited capacity of the state to quickly suppress an insurgency and more likelihood of parity between state and insurgency.

This combination of findings suggest that, rather than any straightforward geographic, economic, political, or social set of determinants, the drivers of intensity have more to do with how a conflict will develop than what incites the original conflict. (In fairness, since most low-intensity conflicts and wars start as small, low-intensity conflict, these findings may also reflect a coding effect of focusing on the most intense phases of a conflict in any given period.)

One of the more surprising findings has been that Polity IV scores have shown a positive correlation with onset of hostilities in general and several sub-types of conflict, including low-intensity conflict. Determining the exact reason why higher Polity IV scores are associated with conflict – in contrast to many other analyses (Collier and Hoeffler 2004; Miguel 2004; Hegre et al. 2001a) – requires going beyond the full-scale Polity IV result. This finding, which is explained in greater detail later, appears to be a function of the impact of mixed regimes, particularly weak democracies, in the data. When the three primary classes of regimes were separated and added to the base regression (using democracy as the base), both autocracy and anocracy showed more correlation with various types of conflict, particularly anocracy. This trend reinforces what many authors have noted, which is that mixed regimes are the most conflict-prone, but it also adds the important element that the highest risk concentration appears to be in democratic-leaning anocracies. For low-intensity conflict, though, there is a slight outlier since low-intensity conflict is least likely to occur in an autocracy, a finding that reinforces the overall thesis that different types of conflict are subject to different correlates.

One caveat to these results: we must consider that, while the regression approach taken for this project is useful for differentiating between governmental and territorial hostilities and between wars and lower intensity conflicts, the approach does not fully indicate whether conflicts start from a common set of causes and then develop into governmental or territorial conflicts based on their circumstances. When onset is recorded, it is marked for the beginning of the highest point of the conflict. Thus, some hostilities could conceivably have begun as lower intensity conflicts with less than 1000 battle-related deaths per year, peaked once over that limit, and then returned to a lower level of intensity. While it was noted earlier that there are only two examples in the data set of transfers between periods, there are 20 cases of periods that mixed – i.e. having both years with fewer than 1000 battle-related deaths and years with 1000 or more battle-related deaths. In these instances, the hostilities would be denoted as a war because the earlier start of the battling as a lower intensity conflict would not be reflected. In general, though, the number of battle-related deaths in war years was much higher than 1000; the averages for the number of deaths in any year with more than 1000 deaths is actually roughly 8500 battle-related deaths. Similarly, the average number of deaths in any period completely classed as lower intensity conflict is substantially below 1000 – much closer to about 330 battle-related deaths per year. Ideally, this study could be taken down to the annual level, which would avoid such complications, but the available data was not well-suited to that pursuit and using five year periods has still produced meaningful results. Thus, it is worth exploring the variables that help differentiate wars from lower intensity conflicts, and these results indicate that the approach of differentiating conflicts based on their intensity has merit and outlines two distinct groups of hostilities.

Impact of Mineral Deposits

The effect of mineral deposits on civil war has been heavily debated, and the extensive data work done for this dataset targeted the issue of different mineral effects. One of the unexpected results is that distance to secondary diamonds is negatively correlated onset of territorial hostilities. Given the strong theoretical arguments in favor of the significance of lightweight, high-value minerals for fueling conflict (Le Billon 2001b, 2005), the result is slightly surprising. Still, previous empirical studies of the relationship between civil war onset and secondary diamond deposits found a positive correlation between the two, but only in a bivariate context (Lujala et al. 2005). In my study, there is also a weakly negative correlation between secondary diamond density and territorial hostilities. When other factors are included in the regressions of previous studies, the results are less definitive, suggesting that the results from this study are not fully counter to the existing literature (Lujala et al. 2005). My study of conflict data does show the fairly robust finding that secondary diamond deposits are negatively correlated with territorial conflict, and that finding is consistent with previous disaggregate studies in Africa (Buhaug and Rød 2006). This finding may be heavily affected by the fact that the major hostilities in which diamonds were involved (Sierra Leone, for example) were classed as governmental conflicts and not territorial conflicts by the Uppsala/PRIO system.

In the regression results for onset, different mineral types have pronounced different effects. Minerals in general show indications of having a positive relationship with onset of governmental war and a negative relationship with territorial hostilities in general. For example, the findings for iron density indicate that a negative relationship with territorial war, and no significance with other conflict sub-types. This finding suggests that the theoretical framework that minerals, such as iron, which have lower value to weight ratios and require more state infrastructure to extract and bring to market will inhibit conflict by increasing infrastructure development and state presence in a local area. While a wide range of mineral types are included in the dataset, iron was used along with secondary diamonds to represent two ends of a spectrum of value-to-weight ratios and ease of access for extraction.

Lag Effects

In all regressions, the spatial and period lag effects are quite strong. While previous aggregate civil war studies have generally had r-squared values of less than 0.3 (Collier and Hoeffler 2004), the regressions for onset of hostilities in this study have r-squared values that are generally quite high, ranging from 0.38 to 0.90. The extremely strong amount of variance explained by the variables included in the regression is the largely the result of the extremely significant and strong lag variables.

One of the challenges in constructing a satisfactory set of lag variables, as discussed earlier, is that some amount of difference from the dependent variable is needed. In the case of spatial lag variables, this situation becomes particularly acute since the spatial lag variable is essentially just a measure of the density of other nearby conflicts. With the most restrictive measures of density, the spatial lag variable collapses towards the dependent variable. As a result, a researcher can also take the reverse approach, intentionally diffusing the density to a great extent in order to increase the overall significance of the other variables. In that sense, the spatial lag variable can become a lever on the significance of the other variables. That said, removing the spatial and period lag variables entirely does not resolve the issue because without a lag variable to help control for the rough patches, fewer of the variables are significant. Without lag variables of any sort, some of the regressions still have strong r-squared values, ranging from 0.21 to 0.68. While it is not a perfect comparison between a regression of many sub-national grid cells and the Collier-Hoeffler study at the aggregate state level, the test of regressions without any lag variables suggests that some of the non-lag variables have meaningful relationships with civil conflict.

In further research, it might be tempting to construct the spatial lag variable after the fact, by intentionally and openly picking the spatial lag density that creates the most significance in the other variables. In other words, it might be an appealing and rational approach to develop the spatial lag variable as dependent on the results of the regression itself, to maximize the explanatory power of the overall regression and create a less arbitrary manner of selecting the kernel radius for the spatial density. It could then be noted what kernel density or other spatial lag variable construction maximized the explanatory significance of the regression overall.

Ultimately, in selecting a lag variable, the lag of the occurrence of conflict – rather than the lag of the onset of conflict – was used. This mirrors the work of other aggregate studies that used “prior war” not “prior onset” to control for the effects of prior battling (Fearon and Laitin 2003b; Collier and Hoeffler 2004). In practice, this controls for the more relevant consideration: that hostilities began because of the spread of conflict across space and time, not because of any factors about the political, social, economic, environmental, or geographic specific to the location. Thus, what is important for the lag variable is the presence of conflict, not the presence of conflict onset. One debatable choice in this use of lag variables is that it uses all hostilities as the basis for the lag *for the onset model only* – rather than different lag variables based on different types of conflict. This formulation suggests that any kind of hostilities, regardless of type or level of intensity, has the potential to impact the likelihood of any other type of hostilities, regardless of the level of intensity. One could argue either that territorial conflicts are actually less likely to spark governmental conflict and more likely to spark other territorial conflicts, as efforts to pull apart a large state engender similar efforts. Also, one could argue that high intensity wars are only likely to spark other high intensity wars and actually discourage low intensity conflict from occurring. These types of arguments, while interesting and the possible topic of future research, are not included in this regression analysis because it would extend the focus of the project beyond its core aims.

Primary Commodity Exports

In each regression the percentage of primary commodities as a share exports and its square are included, in a variable taken from the Greed and Grievance dataset by Collier and Hoeffler (Collier and Hoeffler 2004). In the studies done by Collier and Hoeffler, this pair of variables shows a parabolic relationship between primary commodity exports and civil war, with the most risky values for primary commodity exports being about 33% of GDP. In other words, this parabolic relationship is concave – with a maximum value around 0.33 – but it only applies to the primary commodity exports in general. For my base regressions, none of the curves for primary commodity curves were significant, although a dummy variable to measure whether oil is the dominant primary commodity was negatively correlated with territorial hostilities and low-intensity conflict while being positively correlated with governmental war. As others have noted, the Collier-Hoeffler primary commodity exports variable combines a wide range of different exports – from oil to agriculture (Fearon 2005). Collier and Hoeffler differentiated the primary commodity export variable into different groups, and included a value for when oil is the primary commodity export. Interestingly, the primary commodity export curve for oil is convex, indicating that middle-range values are less risky. Collier and Hoeffler do not directly engage this distinction in their analysis.

For this project, a range of different combinations of variables related to primary commodity exports were tested. Based on the results, it appeared worthwhile to include general primary commodity exports as a percentage of GDP and a binary variable for whether or not oil is the "dominant" commodity, as coded in the Collier-Hoeffler dataset. Collier and Hoeffler note that other disaggregations, for "food, non-food agriculture…and other raw materials" were not significant in their study (and are not included in their dataset) (Collier and Hoeffler 2004). A set of regressions (not shown) with both all primary commodity exports and primary commodity exports only when oil was dominant was conducted, and it did produce improvements in the regressions' explanatory significance as well as producing consistently concave primary commodity curve and consistently convex oil export curves, similar to the Collier-Hoeffler study. The reason for using just the oil dummy variable rather than multiplying the percentage of primary commodity exports as a percentage of GDP by the oil dummy variable is that the oil variable shows signs of having a linear relationship, rather than a parabolic one, because it was not consistently convex or concave in the regressions tested and the arc of the parabolic relationship was small in most cases. Moreover, the net effect of the percentage of primary commodity exports if oil is dominant rarely included more than one change in sign, meaning that the net effect had a single breakpoint between a positive or negative effect on conflict or war. In this context, a linear relationship seemed a more reasonable modeling approach for the oil variable.

In the final regression, the primary commodity exports in general were included in a parabolic relationship along with the dummy variable for oil as the most dominant primary commodity export. These results suggested that oil increases the likelihood of governmental war while decreasing the likelihood of territorial hostilities. The justification for this trend could be that when significant oil exports are present, capturing the state apparatus becomes a compelling enough prize that any violence becomes state-focused. Also, because oil makes the state such a compelling prize while also providing the state with resources to combat an insurgency, oil being the dominant primary commodity export appears to increase the intensity of a conflict as it is associated mostly with governmental war. In the data that is included in the project, the major cases of onset of government hostilities have been the conflict in Angola in the late 1970s and in Algeria and Republic of Congo in the early 1990s. Potential counter-examples for this trend do exist; for example, the rebel activity in the Niger Delta in Nigeria focused on ensuring a measure of autonomy and local control of oil revenues. However, these battles –such as the Oganiland fights of the 1990s – are not classified as civil conflict in the Uppsala database – and the earlier conflicts – such as efforts to seek an autonomous state of Biafra in the late 1960s – fall outside of the scope of the project from 1970 onwards (Osaghae 1995; Gleditsch et al.

2001). Although these potential counter-examples are not present in the data for this project, the basic structural framework still holds: as the value of the state increases and revenues from oil cannot be enjoyed without capturing the entire government, governmental control rather than just territorial secession is the more likely target of armed opposition groups.

Forests

I included forest cover data in a range of different regression models, but I ultimately found the strongest theoretical backing and the greatest significance for two specifications: local forest density and a measure of the difference between local forest density and the state average. Both of these measures exhibited significance in two circumstances. Forest cover density (when substituted for local rainfall, with which it has a high correlation) was negatively associated with territorial hostilities in general and low-intensity conflict in general. In addition, local forest density compared to the state average was positively correlated with territorial hostilities. This contrast – that forest cover in general reduces the likelihood of territorial hostilities but levels above the state average increase the risk – suggests that the territorial conflicts may not be the product of "grievance" from horizontal inequality – such as a lack of resources relative to the capital – but rather by a "greed" dynamic and as areas with comparatively more resources seek to capture that territory.

In regression testing for a possible parabolic relationship between forest cover and conflict, no significant results emerged. When forest level and the forest level-squared were included in the basic regression outlined earlier, only the linear variable was significant. However, in testing a variable that weighted mid-range scores for forest cover highest (forestmidlocal) in a non-parabolic fashion, a strong set of correlations emerged between mid-range scores and governmental hostilities (and its sub-types) and civil war, and a negative relationship with territorial conflict. The counter was also tested: extreme forest values emerged as significantly linked to lower-intensity conflicts. In other words, while the trend is not absolute, in partially forested areas, the hostilities were more intense, while in either very densely or very sparsely forested areas, conflicts were less intense. At a structural level, this finding makes sense – for hostilities to become more intense the environment needs to allow for regular and intense contact between fighting groups. For the same reason that guerrilla strategists favored rough terrain in mountains and forests for protecting their limited number of fighters from stronger forces, intense terrain leads to less intense fighting. In contrast, more conducive terrain leads to more intense fighting in territorial cases. Although the system for scoring forest level was not directly connected to an empirical value for forest density but rather to stratified forest density categories, the forest density that is most conducive to territorial war is slightly higher than the mean level of forest cover on the continent. Meanwhile, the level of forest that is most inhibits low level conflict is roughly three times the mean. This pairing suggests that two separate processes are taking place and that is does not necessarily follow that certain forest levels inhibit conflict *and* promote war. Rather, it appears that above average levels of forest cover inhibit low intensity conflict in general.

It is also worth noting that mid-range values for local forest coverage was another case where the variable was significant but with opposition signs for territorial and governmental hostilities. Moreover the mixed picture of significance based on further sub-type suggests that, as emphasized before, these different sub-types of conflict appear distinct in their correlates and are not easily explained with a single regression.

Another important issue in the discussions of the impact of forest density whether forest density has a primarily economic effect – making development easier – or a technical effect – making conflict harder to execute. Since the mean rainfall levels are generally viewed as a proxy for economic growth at the local level, and since mean rainfall and local forest cover are strongly directly correlated at a level of over 0.8, this combination of correlations suggest that forest cover may have an economic impact as well. Disentangling the economic and the technical effects of forest cover is not straightforward. Given that primary commodity exports at the state level were not directly relevant in either lower intensity conflict in general or territorial war, it suggests that reading forest levels as making the terrain more difficult could hold the greater sway.

Elevation

Elevation range is one of the few variables that appeared significant to conflicts overall and whose sign remained consistent in all cases where the variable was significant. This consistency of higher elevation ranges being linked with conflict demonstrates the value of testing a variable more closely linked with the theoretical argument at stake – that rough terrain promotes conflict risk. Elevation as an absolute measure and through a range of other specifications was less successful at measuring the significance to conflict.

Geographic differences between governmental and territorial hostilities

While the local economic differences of governmental and territorial hostilities have already been discussed, the geographic differences also bear examining. In previous disaggregate studies, one of the primary distinctions between governmental and territorial studies was the distance from the capital (Buhaug and Rød 2006). The regression analysis for this project shows a similar finding – with a strong and significant positive coefficient for distance from the capital and territorial hostilities, and a negative coefficient for distance from the capital and governmental hostilities. At a basic conceptual level, these consistent findings make sense: the further a group is from the government center (the capital) the less likely that the government becomes the primary target of their insurgency. Interestingly, both governmental and territorial hostilities show that proximity to international borders is a significant correlate with onset of conflict. This suggests that international borders play a broader role in the onset of hostilities, which will be examined in greater detail later in this chapter.

The other major difference in the geography of governmental and territorial hostilities is the size of the country. While the size of the country is irrelevant to governmental hostilities, larger counties are significantly more likely to experience territorial war. This finding also makes structural sense: the larger that a country is, the more territorial expanses can have their own sub-groups seeking autonomy. This scale argument for territorial hostilities also plays out in the demographics. While overall population size and local population density are mostly irrelevant for governmental conflict, territorial conflicts are most likely in less populated local areas in less populated states. The finding bolsters the view that it is the far-flung locations – with few people and therefore likely less government attention – that are most likely to seek autonomy from their government. Previous findings on undifferentiated civil war studied at the aggregate level suggested that war was most likely in larger populations and possibly larger states (Collier and Hoeffler 2004; Fearon and Laitin 2003b). In the sub-group of territorial war, the finding that larger populations are related to greater risk of war onset holds, but lower intensity territorial conflicts – and conflicts in general – are associated with lower populations. At a basic level, having fewer people in a country might be seen to be naturally related to having fewer battle-related deaths because fewer people are available to participate in a conflict.

Oil versus minerals (in onset)

Unlike some other variables, oil has a consistent impact at the local level: the onshore density of petroleum deposits is consistently positively associated with hostilities, including general governmental hostilities and war in general. This suggests a fairly straightforward positive association between hostilities and onshore petroleum, particularly governmental hostilities. This consistent impact of petroleum differs from minerals, whose different types show distinct – and sometimes negative – relationships with conflict, particularly territorial conflict.

Minerals, on the other hand, are more differentiated, and their impact at the local level appears to depend primarily on the type of mineral that is in play. While minerals in general show a positive association with governmental war onset and territorial conflict onset, there are specific minerals – such as iron and secondary diamonds – that indicate that they can reduce the likelihood of hostilities. Specifically, iron is negatively associated with territorial war, while secondary diamonds are negatively associated with only territorial hostilities. Explaining the finding that secondary diamonds have no connection to hostilities in general, and only a negative relationship with territorial conflict is not easy. It may be in part because conflicts in Sierra Leone, an oft-cited case study for the relationship between diamonds and conflict, was classified as a governmental conflict, rather than a territorial conflict. It may also be that there are enough different cases of government conflict that is not a significant factor in *onset*, while it plays a larger role in the duration of the conflict after onset. Iron's largely negative relationship with hostilities of all types, with particular significance for territorial conflicts and wars, suggests that the infrastructure development that goes along with the development of heavy metals, such as iron, either reduces a population's interest in secession or increases that state's access to and presence in a territory, making secession less viable. Differentiating between these two explanations is not easy with the available data, but localized information on school enrollment and per capita income might help to resolve the issue. Regrettably, sub-national data is neither reliable nor widespread for school enrollment in Africa and per capita income is in a similar situation. Moreover economic growth under conditions of mineral extraction is less easily understood by proxy through local rainfall levels.

The broader distinction that onshore petroleum density is more closely and positively associated with governmental hostilities while minerals are more often relevant in territorial cases also makes sense. To capture benefits from a petroleum operation, it is almost always necessary to capture the state itself, whereas the impact of a mineral extraction operation is often more localized, increasing its relevance to different types of territorial hostilities.

The role of international borders

With borders playing such as significant role in several of the regressions for onset, it bears examining what factors make borders relevant. Earlier in the paper, three possible explanations were put forth: that borders permit for smuggling to fund conflict, that borders divide common ethnic groups, or that borders provide safe haven through difficult terrain. To test these three explanations, a regression was established with the standard control variables shown in the base regression plus four interaction terms tested alongside each other: border path distance multiplied by the local range in elevation (bordpathrough), the local density of ethnic borders (bordpathethnic), the local density of mineral deposits (bordpathmineral), and the local density of secondary diamonds (bordpathdiasecdens). In this test, the interaction terms that showed the most significance was the border path distance multiplied by the range of the local elevations in the grid cell, which was taken as a proxy for the roughness of the terrain. The border-roughness proxy only showed significance in instances of negative relationships with low-intensity conflict, particularly governmental low-intensity conflict, and territorial war. The local elevation range by itself was significant in other regression tests, but with different signs suggesting that the relationship with international borders is relevant. The general interpretation of this result suggests that borders are conflict-prone primary when they are easily passable (as measured by terrain roughness). In the context of this dataset, the series of territorial conflicts in the

Sahara, an area with almost no elevation change, are a key driver for this finding, but given the extensive case study and anecdotal evidence of rebels using the distant passage routes for transit and trade. An investigation of the data showing the lowest values for the interaction of border path distance and terrain elevation range where onset of conflict occurred confirm that conflicts in Sudan, Mali, and Niger account for many of the grid cells. The second proxy variable that was found to be significant the border-diamonds variable was very significant and had negative relationship with territorial war. This is consistent with the findings for secondary diamond density alone.

When only border-mineral density was included (without comparison to other interaction terms), the borders-minerals term was positively correlated with territorial hostilities, which is in contrast to a negative association to variables on their own. The finding here strongly suggests that minerals may be useful when they can be trafficked across borders. Moreover, combining the other finding from this review of borders – that borders are most useful to insurgents when they are not in rough terrain – suggests that being able to move people and minerals across remote passable borders is has a positive relationship with territorial war.

Political structure

As noted earlier in my paper, the Polity IV score is one of the hardest variables to consolidate into reasonable five-year intervals. Given the challenges of correctly coding, tracking, and consolidating Polity IV scores in a system that makes rational sense and corresponds to the underlying theories of interest, this project tried three different approaches to presenting data on the form of government and its relationship with conflict. For the base regression, which was used as the set of control variables for testing a range of related hypotheses, the Polity IV complete score (with modifications) was used since it is more comparable with previous study. In a separate regression, reproduced at the end of this chapter, the Polity IV *democracy* score only is also used separately. (As noted earlier, when the Polity IV democracy score is used there are five separate variables that are significant but have opposite signs between governmental and territorial hostilities.) The reason for also using the Polity IV democracy score was that its results were largely directly interpretable as the extent to which democratic operations were helpful in reducing the risk of conflict. The reason for including different measures of political structure was demonstrate the differences between the three approaches used for Polity IV scores: single-sided scoring (just democracy scores), full-scale scoring (using the entire -10 to +10 scale), and anocracy scoring (where the score is the absolute value's distance from 0). These results reveal interesting distinctions between the types of conflict studies in their relationship with various forms of government.

Polity IV scores indicate a broad trend towards conflicts being more likely in environments with higher Polity IV scores, i.e. more democratic regimes. Given previous literature, this result by itself is surprising. Polity IV score is the only variable that has the same sign for both governmental and territorial hostilities. This result is surprising and suggests that the complete Polity IV score may not tell the whole story. In fact, when just the Polity IV democracy score is used, a much more consistent and negative relationship with conflict emerges, including government hostilities in general and low-intensity conflict in general. This finding may appear to run contrary to the earlier result with the complete Polity IV score, but in fact, it reveals an underlying risk group – weak democracies. When a specialized variable (p4_weakdemoc_thesis) was used to measure for that specific to weak democracies (distance from Polity IV = 2.5), it was more strongly significant and consistently positive than a simple measure of anocracy (p4_anoc_thesis, measured as distance from Polity IV = 0).

Findings on the risks of mixed regimes are not new to conflict studies, but this thesis's findings place new focus on weak democracies as demonstrating a particularly strong pocket of conflict risk. Moreover, this thesis is the first to connect Polity IV scores to local, geo-referenced conflict, so it is the first time that a more localized lens can be applied to the impact of Polity IV scores. In this light, the Polity IV scores stand out as the most consistent variable in their relationship with conflict, suggesting a rare point of consistency between sub-types of conflict.

Ethnic group borders and religious fractionalization

This thesis is also one of the first projects to use the recently unveiled Geo-Referencing of Ethnic Groups (GREG) project that provided a digital map of the ethnic groupings across the world. In this analysis, ethnic groupings were considered using two metrics – the proximity to ethnic group borders and the density of ethnic group borders. To include a perspective of a statewide metric for fractionalization, the religious fractionalization metric (gg_rf) is also included in a variant regression that is produced below.

To understand the underlying dynamics that appear in the data, it is important to examine the relationship between density and distance of ethnic borders. Proximity to an ethnic border is a sign that territory is located close to an ethnic division, but does not necessarily indicate broader heterogeneity, while ethnic border density is a proxy for a more heterogeneous ethnic environment. This key distinction allows us to distinguish between ethnic groups that border each other as likely to mix and intermingle and those groups with such high border density that regular intermingling and interaction are more likely.

In reviewing the results of the regression analysis, there is strong evidence that the onset of territorial hostilities, including territorial war and territorial low-intensity conflict, have the strongest relationships with the density of ethic configuration borders. While territorial low-intensity conflict is associated with areas of denser ethnic borders, territorial war has less dense ethnic borders. This suggests territorial wars are more likely in areas that are ethnically homogenous. At a conceptual level, this result also makes sense: groups that are ethnically homogenous, but distant from their state capital – by both geographic distance and ethnic division - are more likely to feel less connected to the state and more unified as a separate territory. Thus, these territorial wars onset when the ethnically homogenous group seeks to exit the state, sparking conflict along the borders where populations are most mixed. Moreover, because of the increased homogeneity of the seceding group, it may be easier for the state to enact more aggressive tactics, resulting in more battle deaths. This dynamic can be seen at play in a number of cases – most recently in the troubled border areas between South Sudan, which is predominantly Dinka, and the primarily Arab northern federal Sudan. Also, while this hypothesis was not tested in the data, one could imagine that in heterogeneous areas located close to state capitals, where governmental war is more likely to start, potentially warring groups might already have some role in the government or military that they would seek to expand, rather than giving up that existing place for the uncertain possibility of secession. With territorial conflicts, when distant ethnic groups are less likely to have a role in the government, the strategy of secession may appear to be a more reasonable approach than attempting to capture control of a state.

I also considered whether statewide factors might be a more reasonable method for understanding conflicts over identity. For this analysis, I used the statewide religious fractionalization variable from Collier and Hoeffler's study (Collier and Hoeffler 2004), and when included in the base regression, it significantly reduced the significance of local ethnic border density. Based on the analysis in my regressions, it appears that statewide religious measures of identity divisions may be more salient than local ethnic divisions. One explanation for this finding might be that the overall religious balance, such as between Muslims and Christians in Nigeria, is more important to the state's propensity for organized conflict than local ethnic divisions, which might lead to some communal rioting but fewer organized armed conflicts. As with many variables used in my study, while religious fractionalization showed no significant with general onset, it demonstrated significance for both governmental and territorial conflict with different signs. This distinction – where religious fractionalization is higher with governmental onsets and lower with territorial onsets – highlights yet another case where a single regression analysis for all conflict types is not sufficient, and neglects to bring out some of the significant distinctions between conflict sub-types. The religious fractionalization variable also had impacts on the significance of other variables, beyond just ethnic border density, such as GDP per capita, which becomes significant in the case of territorial hostilities.

Review of Other Hypotheses

Earlier in this project, a range of hypotheses were laid out regarding conflict onset. In addition to examining these broader issues discussed above, each of the individual hypotheses was examined by adding the relevant variable to the base regression listed below. For some variables, there were alternative findings in the onset components of the occurrence models, and, in such cases, the distinction between the pure onset model and the dynamic logit model for occurrence are noted. The purpose of examining such an extensive list of hypotheses is both to take advantage of such a large dataset to explore a variety of topics and to catalog the different areas of distinction between conflicts based on the actors' aims and intensities. The previous sections have demonstrated some of the significant differences between conflicts based on the actors' aims and the intensity, and this list provides a few more detailed distinctions between the categories.

The table below summarizes the variables involved and the findings that resulted from this extensive series of tests. It is definitely possible that these hypotheses could be tested under different conditions (other than using the base regression as the control set of variables) and receive different results. But, when tested under the specified conditions, some individual results help to fill out the importance of distinguishing between types of conflict. (In some cases, when co-linearity between different data specifications became an issue, the variable was substituted for another relevant variable, such as in different data specifications for the forest data.)

HYPOTHESIS	**RESULTS**
Economic Conditions	
Gross Domestic Product (GDP) and Economic Performance	
Hypothesis 1. Low country GDP (`GDPtotalstate`) corresponds with higher local conflict risk.	Total state GDP has a strong correlation with conflict risk in general, which may provide general support to the "prize" theory of conflict. GDP showed significance for territorial hostilities, low-intensity conflict in general, and civil war in general.
Hypothesis 2. Low GDP per capita (`gg_lngdp_`) corresponds with higher local conflict risk.	Low GDP per capita appears to have a limited relationship with onset in the onset logit model, showing only positive correlation with territorial conflict. In the onset component of occurrence models, the data is more compelling, showing a consistent relationship between low GDP per capita and higher local conflict risk for both governmental and territorial onsets. In the case of territorial hostilities, especially territorial war, low GDP per capita appears to be correlated with both high onset risk and higher risk of continuance of the conflict.
Hypothesis 3. Low country GDP growth (`gg_gy1`) corresponds with higher local conflict risk.	Low country GDP growth corresponds with higher risk of governmental hostilities, particularly governmental war, and territorial low-intensity conflict, but high country GDP growth corresponds with increased risk of territorial hostilities, particularly territorial war. These results are mostly confirmed by the dynamic logit occurrence model, although territorial hostilities onset in general shows a correlation with low GDP growth rather than higher growth in the occurrence model.

Hypothesis 4. Increases in rainfall overall (`raingrowlocal`) correspond with reduced risk of conflict.	Increases in rainfall growth are associated with lower risk of conflict in general and governmental hostilities of all types. This strong finding bolsters the argument that growth, particularly local growth, can stave off fights against the government. In contrast, there is strong evidence that increases in rainfall overall correspond with <u>increased</u> risk for local territorial conflict, with particularly strong effects on local territorial war. (The territorial onsets are largely driven by onsets corresponding with rain increases in Niger, Mali, Ethiopia, and Morocco.)
Hypothesis 5. Increases in rainfall per capita (`raindiffpclocal`) correspond with reduced risk of conflict.	There is evidence that increases in rainfall per capita also reduces general conflict risk, particularly government conflict risk. In contrast, consistent with above, increases in rainfall per capita also are correlated with increased territorial conflict risk.

Horizontal Inequality and Income Distribution	
Hypothesis 6. Countries with higher rates of income inequality (`gg_ygini`) experience a higher risk of conflict onset and occurrence throughout their territory.	There is very limited evidence that rates of country-wide income inequality correspond with conflict in general. Statewide income inequality was significant only when positively associated with general territorial hostilities and negatively associated with general low intensity conflict.
Hypothesis 7. Countries with higher rates of land inequality (`gg_lgini`) experience a higher risk of conflict onset and occurrence throughout their territory.	Land inequality is associated with a higher risk of low-intensity territorial conflict and a lower risk of territorial war. This finding may indicate that high land inequality in a country creates conditions where the government is particularly aggressive about ensuring territorial integrity.
Hypothesis 8. Areas with more than the state average for rainfall (`meanrainlocalVSstate`) have a greater risk of governmental conflict.	There is no evidence of a significant relationship between relative rainfall and governmental conflicts.

Hypothesis 9. Areas with less than the state average for rainfall (`meanrainlocalVSstate`) have a greater risk for territorial conflict.	Similar to having above-state-average levels of forestation, above-state-average levels of local rainfall are positively correlated to territorial hostilities of all types. Surprisingly, this hypothesis does not appear to have significant evidence to support it in the regression analysis.

Rough Terrain

Hills and Mountains	
Hypothesis 10. Higher elevation (`elevlocal`) corresponds with higher conflict risk.	Absolute elevation levels show no particular association with conflict.
Hypothesis 11. Changes in elevation (`elevrangelocal`) correspond with higher conflict risk.	Elevation changes show a consistent positive correlation with higher conflict risk in general, particularly for governmental low-intensity conflicts and governmental civil wars. The significance for governmental hostilities is confirmed in onset component of the dynamic logit occurrence model, and territorial hostilities also appear to have a positive correlation with conflict risk in the occurrence model.
Hypothesis 12. High slope gradients (`elevslopepctlocal`) correspond with higher conflict risk.	There is significant evidence that higher slope gradients correspond to higher conflict risk, particularly for government conflicts. Higher local slope gradients were also correlated with higher likelihood of both lower intensity conflicts and civil war. This finding is generally consistent with the finding for changes in elevation (`elevrangelocal`).
Hypothesis 13. Hills – mid-level ranges for the elevation (`elevationhills`), change in elevation (`elevrangehills`), and slope gradient (`elevslopepcthills`) – correspond with higher risks of conflict in a parabolic relationship, while very mountainous and very flat areas are less conflict-prone.	In the regression analysis, mid-ranges for elevation and mid-ranges for changes in elevation were not significantly related to hostilities in general, but did have significance for territorial conflict (with a significant negative association) and territorial war (with a significant positive association).

Foliage and Forests

Hypothesis 14. Increased forest density (`forestcoverlocal`) is associated with higher risk of conflict.	There is no evidence that increased forest density is associated with higher risk of conflict. To the contrary, increased forest density appears to reduce the likelihood of territorial hostilities of all types.
Hypothesis 15. There is a parabolic relationship between forests and conflict, with both densely forested and sparsely forested lands associated with a higher risk of conflict (`forestmidlocal`).	There is little evidence of a parabolic relationship that favors extreme conditions for conflict. Rather, there is evidence that middle values are more likely to lead to governmental hostilities in general and civil war in general. Territorial hostilities appear to be more common in extreme circumstances, although that may be a function of the data for this project featuring several significant territorial conflicts in the Sahara.

Natural Resources

Oil

Hypothesis 16. Proximity to onshore oilfields (`petrodist`) corresponds with increased risk of conflict.	Distance to onshore petroleum has no significant relationship with governmental conflict. On the other hand, as distance from petroleum resources increases, the risk of territorial hostilities, and territorial war in particular, decreases significantly.
Hypothesis 17. Density of onshore oilfields (`petrodense`) corresponds with increased risk of conflict.	Evidence indicates a significant association between onshore petroleum density and armed hostilities, particularly governmental war. There is a particularly strong positive relationship between onshore petroleum density and governmental hostilities as well as higher intensity conflicts (wars). This suggests that this hypothesis is confirmed for local conflict.

Hypothesis 18. Onshore oil deposits (`petrodense`) are likely to increase the number of battle-related deaths from a conflict area.	Evidence indicates that onshore petroleum density is significantly positively related to war, with has more than 1000 annual battle-related deaths. Since there was no positive association with lower intensity conflict, there is evidence that onshore oil deposits are associated with more intense hostilities.
Hypothesis 19. Onshore oil deposits are likely to decrease conflict when they are located near to the state capital and increase conflict when they are located further from the state capital (`cap_petro_proximity`).	The terms of specification for this hypothesis are relevant to the outcome. If the measure uses an interaction term between proximity to the state capital and *density* of onshore petroleum deposits (petrodense), then the hypothesis holds that government war is less likely in areas where petroleum is located near state capitals. If the specification uses an interaction term between distance to the state capital and the distance to an onshore petroleum deposit (petrodist) then territorial conflict and territorial war are more likely to occur in those regions, with no clear significance for governmental conflicts. This suggests that territorial conflict may be driven in part by the "greed" explanation, such as an opportunity to capture outlying oil deposits.

Diamonds and other Minerals

Hypothesis 20. Primary diamond deposits (`dianldense`) inhibit both conflict onset and conflict occurrence.	Primary diamond deposits appear to decrease the likelihood of conflict in general. Notably, because there is no territorial conflict in places with a kernel density of primary diamond deposits are greater than 0, the dianldense variable had to be removed for collinearity.

Hypothesis 21. Secondary diamonds (`diasecdistln`) will make areas more conflict-prone to both onset and occurrence.	There is little evidence that secondary diamonds make conflict more likely in general. In instances of territorial hostilities, regardless of intensity, conflict may be less likely to onset in areas closer to secondary diamonds. Moreover, territorial conflicts in general may have a longer duration in areas that are farther from secondary diamonds. This result partially runs counter to existing theory that low-weight, high-value minerals are more “lootable” and increase conflict risk. This result may be because the most well-known cases of conflict diamonds, such as Sierra Leone, are classed as governmental conflicts. The notable exception here is that territorial wars are likely to last longer in places with closer proximity to secondary diamonds. This may reflect that secondary diamonds increase intensity as well as duration in extreme cases, consistent with some case study evidence (Weinstein 2007).
Hypothesis 22. Diamonds located closer to the state capital (`diasecdens_cap`) inhibit conflict, but that diamonds located closer to a state border (`diasecdens_bord`) will make conflict more likely in those areas.	While the relationship between secondary diamond density relative to state capitals had no relevance for any subtypes of conflict, the interaction term between secondary diamond density and international borders made territorial conflict, specifically territorial war, more likely.

Hypothesis 23. Gold (`golddense`), copper (`copperdense`), platinum (`platinumdense`), manganese (`manganesedense`) and coltan (`coltandense`) correspond with higher risk of onset and occurrence of conflict, while iron (`irondense`) is negatively associated with the onset and occurrence of conflict.	Despite extensive regression analysis, the findings on minerals are generally mixed with a slight tendency for minerals to reduce conflict, particularly territorial war. Gold and copper deposits strongly decrease the likelihood of governmental war. Manganese deposits show the most significance across different types of conflict; manganese had negative correlations with territorial conflicts of all types but positive correlations with low intensity conflict, particularly governmental. These many splits reinforce not only the challenge of using variables with limited onsets and the limited mineral deposits but also the difficulty of accurately modeling all types of conflict with a single regression. Platinum deposits correlated with the likelihood of low-intensity conflict, particularly governmental, but to significantly decrease the risk for territorial war. Coltan deposits are negatively associated with risk of territorial low-intensity conflict. Iron had no significance with conflict in general but appeared negatively correlated with territorial war and governmental low-intensity conflict.
Hypothesis 24. The presence of mineral deposits (`mineraldens`), regardless of their type, is increases the number of battle-related deaths.	The evidence is not completely clear on this issue, and the effect appears to be related to the type of conflict. Mineral deposits in general are associated with governmental war, but not lower intensity governmental conflict. However, mineral deposits are negatively associated with lower intensity territorial conflict in general.

Population

State Population Size, Density, Growth and Distribution

Hypothesis 25. Higher state populations (`poptotalstate`) are associated with increased likelihood of local conflict.	Total state population appears to have no impact on the risk of governmental hostilities. For territorial hostilities, total state population appears to make conflict more likely to be severe, as total state population is positively associated with territorial war and negatively associated with lower intensity conflict in general.
Hypothesis 26. Higher state populations (`poptotalstate`) make conflict more likely in the areas far from state capitals.	Since territorial war is more likely to start far from a state capital and total state population is positively associated with territorial war, it appears that higher state populations do partially make conflict more likely to occur far from state capitals.

Local Population Size and Density

Hypothesis 27. Areas that experience significant population increases (`popratelocal`) have a higher risk of civil conflict.	Local population growth rates do not appear to influence conflict risk in general. There is slight evidence that local population growth makes governmental war less likely.

Military Capability

Hypothesis 28. States with high military expenditures (`cow_milex`) are less likely to experience violent civil conflict anywhere in their territory.	States with high military expenditures experience less risk of territorial hostilities and civil war. While it cannot be said that high military expenditures reduce the risk of all types of conflict, it appears that they reduce the risk of several sub-types of conflict. This finding suggests that adequate military spending can help to reduce conflict.

Hypothesis 29. States with high military personnel per capita (`cow_milper_pcstate`) are less likely to experience violent civil conflict anywhere in their territory.	States with higher military personnel (cow_milper) show a slightly higher risk of conflict in general, particularly territorial hostilities in general and government war. States with higher military personnel per capita (cow_milper_pcstate) confirm those results, showing stronger likelihood of territorial hostilities, particularly territorial war, and low-intensity conflict in general. This finding, coupled with the finding on military expenditures, supports the view that large numbers of military personnel can pose a risk to the unity of the government. (Note: Regressions for conflict in general and government hostilities onset did not converge for cow_milper_pcstate.)

Ethnic Fractionalization

Hypothesis 30. Conflict onset and occurrence will be more likely in areas with high density of ethnic borders (`greg_densnocoast`).	High densities of ethnic borders is positively associated with territorial conflict in general, particularly lower intensity territorial conflict but negatively associated with territorial war. This finding suggests that minor battles to control territory are more likely in ethnically divided areas while severe battles to secede from a state are more likely in ethnically homogenous local areas.
Hypothesis 31. Conflict onset and occurrence will be more likely near ethnic borders (`greg_distnocoast1`).	There is strong evidence that territorial hostilities of all types are more likely to onset far from an ethnic configuration border. In contrast, governmental low-intensity conflict is more likely nearer to an ethnic configuration border. This suggests that battles to secede are more common in places further from ethnic groupings mixing.

Country Size

Proximity to the Capital

Hypothesis 32. Territorial conflict is more likely in areas further from the state capital (`cappathdist`).	There strong evidence that territorial hostilities (as well as both territorial war and lower intensity territorial conflict) occur further from the state capital. These findings are confirmed in the dynamic logit occurrence model.
Hypothesis 33. Governmental conflict onset is more likely in areas closer to the state capital than territorial conflict onsets.	While there is some evidence that territorial conflicts are more likely to occur further from the state capital, there is limited evidence that governmental conflicts onset closer to the state capital than territorial conflicts. Rather, governmental conflicts tend to be more driven by relationship to borders, a variable that is only partially correlated with distance from capitals.

Proximity to an International Border

Hypothesis 34. Conflict is more likely in border areas where there is a high concentration of secondary diamonds (`bordpathdiasecdens`).	Conflict in general does not have any relationship with secondary diamonds along borders, although there is a strong negative correlation between secondary diamond concentration along borders and lower risk of territorial war. This finding appears surprising based on existing theory (Le Billon 2005), but is not inconsistent with other studies (Lujala et al. 2005).
Hypothesis 35. Conflict is more likely in areas that have ethnic group configurations whose territory crosses an international border (`bordpathethnic`).	There is no evidence of a general relationship between ethnic group divisions near international borders and conflict, except for a higher risk of low-intensity territorial conflict.
Hypothesis 36. Conflict should be more likely in border areas that have rough terrain (`bordpathrough`).	The border-roughness term appeared to reduce the risk of low-intensity conflict in general, particularly governmental conflict, while increasing the risks for territorial war. Identifying the operational cause of the border-roughness is not easy: one explanation, which would be consistent with the findings for border-ethnicity, is that the risk for territorial war may be increased because outlying groups are further separated from the center of government by elevation changes.

Hypothesis 37. The capital-border distance (`capborddist`) difference increases so that locations near capitals and far from borders are least likely to have conflict while those near borders and far from capitals are most likely to have conflict.	Similar to simple distance from the capital, only with a slightly stronger effect, the capital-border distance significantly increased the risk of territorial hostilities of all types.

Form of Government – Democracy Anocracy, and Autocracy

Hypothesis 38. States with anocratic regimes will be more likely to experience conflict onset and occurrence in their territory.	Anocracy – middling Polity IV scores – is associated with higher risk for conflict in general, governmental hostilities, and low-intensity conflict in general. This finding is consistent with other studies and, upon further examination, appears most significant in the weakly democratic regimes that center on Polity IV scores of approximately 2.5.
Hypothesis 39. Autocratic and democratic states will be less likely to experience conflict onset and occurrence in their territory.	In the converse of anocracy, strongly democratic or autocratic regimes appear less likely less likely to experience conflict in general, although there is evidence to suggest that slightly higher Polity IV scores pose a greater conflict risk in general.
Hypothesis 40. The proximity to the capital and the political regime type have an interaction effect that makes conflict more likely near the capital in autocratic regimes (`cappathdist_autocracy`).	Proximity to the capital and strong autocracy have a strong negative relationship with the likelihood of conflict in general, particularly government hostilities. Proximity to autocratic capitals was negative correlation with low-intensity conflict and civil war as well. This finding is strong evidence for the local stabilizing effect of autocracy around a capital.
Hypothesis 41. Territorial and governmental conflicts are more likely to occur far from the state capital in democratic regimes (`cappathdist_democracy`).	Conflicts did not appear to be appreciably more likely in locations further from the capital in democratic regimes. There is not significant evidence to support this hypothesis.

Base Regression for Onset

Variable	Base General Conflict Onset	Base Governmental Hostilities Onset	Base Territorial Hostilities Onset	Base Low-Intensity Conflict Onset	Base Civil War Onset	Base Governmental Low-Intensity Conflict Onset	Base Governmental Civil War Onset	Base Territorial Low-Intensity Conflict Onset	Base Territorial Civil War Onset
Spatial Lag	1.5522745***	1.0707547***	2.3222629***	1.3370114***	1.1187107***	1.0575436***	.92405858***	2.2728658***	3.1210658***
Spatial Period Lag	-.83414489***	-.34305749*	-1.7789383***	-.73583533***	-.39523591*	-.48732781**	-0.06275995	-1.3392701***	-4.5161123***
Country size	-0.00079828	-0.00103688	0.00022118	-0.00152067	0.00002814	-0.00054868	-0.00053037	-0.0031916	.00451322***
Mean local rainfall	-0.00035864	-0.00015963	-.00190384***	-0.00059352	-0.00013642	0.00015473	-0.00022119	-.00249538***	-.00185288**
Local forest cover vs state	-0.06835298	-0.16669523	1.1961594**	-0.00507508	-0.03668375	-0.16195042	-0.17066596	1.0404073**	0.10658342
Change in local rainfall per	-.00579143**	-.00782597***	.00707258*	-0.00460143	-.00580974*	-.00821322***	-0.00503629	0.00515139	0.00789443
Local elevation range	.00070405*	.00136148***	-0.00104577	.00078029*	0.00066504	.00103824***	.00142759***	-0.00026286	-0.00187296
Local population	.11436699**	.17825066***	-.19149028***	0.09781522	.11642827**	.24485203***	.13395752**	-.20020976**	-.11648904**
Total state population	-0.000003002	-0.000003028	-0.00000103	-7.451e-06*	-0.00000249	-0.00000941	-0.000003611	-0.000006382	.00003885***
Mineral deposity density	0.00894329	0.02298583	-.59495982**	-0.00351396	0.09083351	-0.03693761	.12879817*	-0.46492641	-0.52858795
Distance to sec. diamonds (nat.	-0.02224826	-0.21871279	1.3932146***	0.46220014	-0.11838157	0.14568479	-0.20116888	1.9767617***	1.5480438***
Density of onshore oil	.36876394**	.29629452**	1.3362022	0.56644525	.30206144**	0.4307175	.30838968**	0.68324179	0.9364764
Path distance to state capital	5.372E-07	1.263E-08	1.996e-06***	4.868E-07	2.151E-07	-3.202E-07	4.864E-08	1.088e-06*	2.381e-06***
Proximity to state borders	2.350e-06***	3.089e-06***	-5.155E-07	9.923E-07	3.582e-06***	1.801e-06**	3.446e-06***	-0.000001638	9.461e-06**
State military expenditure	-0.000000129	4.328E-07	-5.705e-06***	6.361E-07	-9.230e-07*	1.717e-06**	-1.106e-06*	-.00001312***	-6.454e-06*
State military personnel	.00797375*	0.00287423	.02876767***	0.01054687	0.00898548	-0.0018762	.01017774*	.06171472***	-0.00679519
GDP per capita (natural log)	-0.17218974	-0.2602166	1.5684885	0.25404522	-0.22634763	0.25153617	-0.38004524	.9450147***	2.570675
GDP growth (GG)	-.19234728**	-.32414994***	.34788911***	-0.04971799	-.22718908*	0.02807871	-.45567403***	-.29460843**	1.6035876***
Primary commondity	3.3660301	2.3071932	32.454135	-7.2823918	5.6124753	-5.3715903	11.076316	-2.2128812	11.966785
Primary commondity	-3.3888488	-7.6945697	-191.86283	16.280587	-10.139776	9.385094	-25.955857	-45.097168	-45.906792
Oil primary export (dummy)	-0.34806244	1.0778215	-8.0724132***	-4.1149197*	1.0539513	-2.8888007	2.8685225***	-0.70879176	
Polity IV score (thesis	.14240072***	.11932923**	.19348358**	.15533097**	0.06862461	-0.03594364	.19272891***	.6704225***	-.83428372***
Density of ethnic borders	-0.01135748	-0.06079392	.55088329**	-0.00569489	0.01699232	-0.20368531	-0.01503735	.92948061***	-.3481611*
Constant	-5.3209913**	-3.025188	-36.904777***	-12.27822*	-6.0741843**	-10.343813*	-4.4620848	-36.519412***	-73.216776***
R-squared (adj.)	0.54461024	0.48754912	0.79220516	0.45456541	0.47911609	0.36504185	0.55179237	0.79939735	0.90630008
Grid cells	83587	83587	83587	83587	83587	83587	83587	83587	63877

* p<.1; ** p<.05; *** p<.01

Conclusions on Onset

At the beginning of this paper and then again at the outset of this chapter, I posed the fundamental research question of whether conflicts are worth distinguishing by the aims of the conflict actors or the intensity of the fighting. Through an exhaustive review of hypotheses and a comparison of regression results, it is apparent that conflicts that seek territorial secession have substantially different correlates than those that seek government capture. Moreover, there are significant distinctions in factors that correlate with occurrence between low-intensity civil conflict and higher intensity civil war.

At a broad analytical level, what emerges from distinguishing conflicts based on their aims and intensities is that a single analytical approach or regression is unlikely to accurately describe all conflict types. In some cases, such as GDP growth and local population density, the overall findings may even mask a significant counter-finding for an important subgroup of conflicts, such as governmental or territorial hostilities. Moreover, the hybrid approach taken by this project has allowed aggregate and disaggregate variables to be considered together over both time and space. One of the analytical outputs of that process is that statewide variables economic status – GDP per capita – appears less significant to local conflict onsets than originally expected based on other aggregate studies using similar statistical approaches. Without taking a hybrid approach that included statewide variables measured for changes across time, and local geographic / environmental factors that change little over time but much within countries, this finding would have been difficult to conclude or effectively compare. This finding further suggests that as conflict studies become more interested in low-intensity conflicts, adopting a disaggregate approach to their study may improve the quality of results for future research.

Determining the correlates for the onset of different types of conflicts is not necessarily the same as identifying the pressure points that could reduce the length of the conflict, as the following section will demonstrate, but it does suggest that the identification of this set of correlates associated with the start of conflict could be useful in targeting conflict prevention activities. Moreover, the advantage of this hybrid approach is that it highlights opportunities to engage at the statewide political and economic level, as well as the local environmental level, to help reduce conflict risks.

Occurrence of Conflict

Note: In past studies, conflict "occurrence" has also been termed conflict "incidence" or conflict "prevalence." These concepts are all generally consistent and refer to the presence of ongoing conflict, not merely the first year of conflict onset.

For this project, given the depth of the dataset, exploring methods for modeling the occurrence –

Totals n = 95,886 inc. 19419 **occur** (20.25% occur)	Governmental Hostilities n = 12900 cells	Territorial Hostilities n = 6769 cells
Civil War n = 13423 cells	Governmental War n = 9495 cells	Territorial War n = 4168 cells
Lower-Intensity Conflict n = 6006 cells	Governmental Conflict n = 3405 cells	Territorial Conflict n = 2601 cells

rather than just the onset – of conflict was also attractive. While the approach employed in this chapter instills less confidence given the challenges of modeling for spatial and temporal lags, it does match with the approach used by others working in the field of modeling civil war (Fearon and Laitin 2003a). As described earlier in the Methodology section, the method used is termed a “dynamic logit” by Fearon and Laitin, who use it to conduct some modeling for civil war occurrence based on their worldwide country-year dataset. This modeling approach calls for including all of the terms of the regression in an interaction with a lag of the dependent variable. The concept behind this approach is to separate the portion of the regression that represents onset (the variables themselves) and the portion that represents non-onset occurrence (the interaction terms). Since the dichotomous lag variable reflects a 0 when onset occurred and a 1 for the remainder of the conflict, multiplying all of the coefficients separates the coefficients that represent the onset (the first period of the conflict) and those the interaction term coefficients that represent the remainder of the conflict (its continuance past the onset). Combined, these two sets of coefficients – one from the variables themselves representing the onset and another from the interaction terms representing the remainder of the conflict – can be used to assess both the likelihood and the potential duration of a conflict.

For Fearon and Laitin, this approach generates results that are mathematically pleasing and conceptually compelling because the initial model for onset (without any interaction terms) and the onset portion of the combined dynamic logit look very similar (see table 1 from “Additional Tables for ‘Ethnicity, Insurgency, and Civil War,’ ” page 12). The similarity between the “effect on onset” and the “base model” for onset lend credence to the modeling approach and reinforce robustness of the base model.

Table 1: Different Estimation Methods

	(1) Base model	(2) Dynamic logit	
Dep. Var.	onset	war	
		Effect on onset	Interaction with lagged war
Prior War	-0.954**		6.641**
	(0.314)		(1.498)
Per Capita Income (lagged)	-0.344**	-0.362**	0.472**
	(0.072)	(0.076)	(0.131)
log(Population) (lagged)	0.263**	0.253**	-0.056
	(0.073)	(0.080)	(0.143)
log(%Mountains)	0.219**	0.228**	-0.186
	(0.085)	(0.088)	(0.145)
Non-contiguous State	0.443	0.433	0.220
	(0.274)	(0.306)	(0.520)
Oil	0.858**	0.850**	-1.962**
	(0.279)	(0.307)	(0.547)
New State	1.709**	1.755**	
	(0.339)	(0.341)	
Political Instability	0.618**	0.607*	-1.318**
	(0.235)	(0.260)	(0.385)
Democracy (Polity 4)[a]	0.021	0.027	-0.041
	(0.017)	(0.018)	(0.034)
Ethnic Fractionalization	0.166	0.111	1.204
	(0.373)	(0.398)	(0.659)
Relig. Fractionalization	0.285	0.343	-0.444
	(0.509)	(0.542)	(0.883)
Constant	-6.731**	-6.616**	
	(0.736)	(0.784)	
Observations	6327	6327	

Figure 21. Reproduced from Fearon and Laitin's "Additional Tables on Ethnicity, Insurency, and Civil War."

In the case of Fearon and Laitin, the total explanatory power of the regression was limited both by having less data and by having no strong spatial lag effects. While Fearon and Laitin tested a variable for civil wars in neighboring countries in the prior year as part of their regression analysis, it did not return as significant, possibly because civil war in neighboring countries appeared in 44% of the observations in their dataset (Fearon and Laitin 2003b). In the case of this analysis for civil conflicts in Africa, the strong effects of the spatial and temporal lags are increased by including them as interaction terms as well, and, as a result, the "effect on onset" portion of the dynamic logit for this project does not appear as similar as for Fearon and Laitin's model. Nonetheless, it is clear that the underlying data and trends between the base model for onset and the effect on onset in the dynamic logit model are similar, and those areas of greatest similarity suggest particularly strong robustness.

Past studies on the incidence of conflict have almost all been conducted on the aggregate level. The most recent literature on conflict incidence, such as recent studies by Besley and Persson, uses mostly country-year approaches along with economic models to determine the conditions under which civil war is most likely. As with most aggregate studies, studies of civil war do not differentiate conflicts based on the aims of the conflict actors, and no conclusions are drawn for low-intensity conflict (Besley and Persson 2009). Earlier studies by Collier and Hoeffler as well as by El Badawi and Sambanis also focused on a mostly economic analysis of the likelihood of civil war (Collier and Hoeffler 2002; Elbadawi and Sambanis 2002). Although both studies were also set at the aggregate level, Collier and Hoeffler used a five-year period approach, but Elbadawi and Sambanis took a country-year format. There has not been as much development of conflict incidence investigation using disaggregate techniques, possibly because of the difficulty of tracking changes over time in a disaggregate format.

The conclusions of studies on conflict incidence have been varied in their focus. For Besley and Persson, the country-year format works well because their most novel findings focus on the impact of global commodity prices on the incidence of conflict (Besley and Persson 2009). In studying the incidence of civil war, Collier and Hoeffler note that the incidence of conflict in Africa is not anomalous based on the continent alone – rather, the conflict pattern in Africa reflects poor economic conditions that are not offset by less conflict-prone societal conditions (Collier and Hoeffler 2002). Meanwhile, Elbadawi and Sambanis determine that political variables (specifically Polity IV scores) and economic variables (specifically real GDP) have particular significance for reducing conflict prevalence when compared with social factors, such ethno-linguistic fractionalization (Elbadawi and Sambanis 2002).

The dynamic logit analysis conducted for this project broadly suggests similar core trends as those identified earlier in the analysis for onset in this project. A few of the more robust trends that were present in the original onset analysis and are maintained in the analysis conducted for the onset effect on occurrence. In one notable shift, GDP per capita appears to be more significant, and with the expected negative correlation with conflict risk, in the onset portions of the dynamic logit models. Also, GDP growth has a significant and negative relationship with territorial hostilities, in contrast to its significant and positive relationship in the logit regression with clustered errors. This shift may reflect some of the changes in lags between the logit regression with clustered errors used for onset analysis and the dynamic logit model used for occurrence analysis. Moving beyond the common findings in onset, there are several findings on occurrence that are worth highlighting from this analysis. Consistent with the general findings of Elbadawi and Sambanis for civil war, the regression analysis for the continuance of low-intensity conflict or civil war finds that higher Polity IV scores reduce the chances that low-intensity conflict will continue. These findings suggest that perhaps in more democratic states, conflicts are more easily resolved either through changes in government policy or political assimilation of insurgents. In contrast, in more autocratic regimes, government may find it difficult to resolve internal disputes because it is unwilling to bend policy based on insurgent complaints. This finding in relation to the impact of political system on the occurrence of a conflict is just one area where taking a disaggregate approach to studying conflict occurrence has produced compelling results. The following sections outline additional areas.

Note: Due a failure of the regression to converge, local forest cover was substituted for local rainfall in the regressions for governmental low-intensity conflict and governmental war. The substitute has an approximately r = 0.80 correlation with the original.

Statewide Economics and Occurrence of Conflict

Statewide economic growth (`gg_gy1`) appears to have little impact on the likelihood of governmental or territorial hostilities continuing in general. In contrast, statewide economic growth decreases the risk of territorial wars continuing. This finding, which is not consistent with the earlier determinations about hostilities in general, may indicate that strong states have more success in limiting the size of territorial hostilities and preventing their growth into wars. While this result for territorial war, versus a non-result for other conflict sub-types, demonstrates the value of segmenting sub-types of conflict in analysis.

A second interesting finding is that while statewide GDP per capita (`gg_lngdp_`) has the potential to extend low-intensity conflicts in general but shorten territorial wars. It may be that higher GDP per capita means that the state and the populace are more willing to tolerate low-level conflicts without resolving them with overwhelming political or military force, which might harm economic prospects. However, when the conflict reaches the level of a territorial war, the willingness to allow the conflict to continue ceases, and the government becomes more active in suppressing the insurgency.

In the case of territorial hostilities, the combined onset and occurrence effects suggest that in countries that are growing economically, territorial wars will be shorter. This suggests that the concept that some countries can grow their way out of conflict may hold for cases of territorial wars, possibly because the economic growth enables the government to sustain more successful military operations or because it reduces the incentives for secession. Parsing these potential twin drivers of the finding is not clear and cannot easily be accomplished using proxy variables in regressions. Given that both high statewide GDP and GDP growth per capita can shorten territorial wars, it may be that both effects are in play with the high state GDP levels more indicative of state capacity while GDP growth per capita becomes more indicative of opportunity cost for insurgents. In the case of governmental low-intensity conflict, higher statewide GDP per capita levels also reduces the duration of conflicts, although GDP growth per capita is not significant in conflict duration. This suggests that, while changes in GDP growth that may be significant in sparking conflict, it is GDP per capita that determines the length of low-intensity government conflicts. For governments, the willingness to tolerate even low-level conflict may be limited if the state has the resources (with GDP per capita levels standing in as a proxy indicator) to end the conflict quickly.

Moreover, at the local level, having more forests than the state average or having an increase in local rainfall can shorten a territorial war. The absence of forest cover – associated closely with limited rainfall, which some analysts interpret as a proxy for local economic growth (Miguel 2004) – is significantly associated with the duration of territorial war. This narrative matches partially with some of the conflicts that have developed around the Tuareg rebellions in northern Mali and Niger, as well as some of the secessionist fights in eastern Ethiopia, where a lack of economic growth coupled with poor local ecology has helped to prolong conflicts.

Political Structure

The most consistent finding in the occurrence data is that higher Polity IV scores lead to shorter conflicts in all cases, even when a higher Polity IV score identified as a possible source of conflict onset. This finding was significant in cases of low-intensity conflicts, civil wars, territorial war, low-intensity government conflict, and governmental war. Thus, while democratic structures may not be the best at preventing conflict, particularly given the earlier findings on the risks of weak democracies, higher Polity IV scores suggest that more democratic regimes may be able to transform opposition groups into political, rather than armed, groups more easily.

Population

Larger state populations tend to lengthen duration in cases of low-intensity conflict, particularly governmental low-intensity conflict. This result may be because larger populations can more easily absorb a small conflict without requiring substantial response. It may also be because smaller insurgencies find it easier to hide in larger populations, where the size of the population makes it a challenge for the state to identify all insurgency members. It is notable that larger state populations generally make low-intensity governmental conflicts more likely, either because the large state prevents minor conflicts or because conflicts grow quickly to wars when they emerge. These twin dynamics are difficult to differentiate, but the dynamic logit approach does helpfully distinguish between causes of onset and continuance, which may be opposing in some cases.

Minerals and Occurrence of Conflict

First, overall mineral density shows an association with shorter duration conflicts while secondary diamond density shows a strong impact in extending the length of territorial hostilities. In the case of territorial war – the effects are reversed: mineral density in general is associated with longer territorial wars while secondary diamond density is associated with shorter durations. In the case of secondary diamond distance, the impact of diamonds extending territorial conflict closely matches existing theory. Given the combination of onset and occurrence data, my study lends partial support to earlier studies that suggested minerals extend conflicts, particularly territorial ones (Malaquias 2001; Ross 2002; Snyder and Bhavnani 2005; Ross 2006), as noted in earlier chapters. Moreover, this finding clarifies what most theoretical discussions of mineral wealth already mention: that the relevance of minerals is not in creating the initial grievance or opportunity that leads to an insurgency, but rather in creating a revenue stream that can be controlled by insurgents or government forces to sustain fighting forces. Thus, it makes sense that minerals should show a much stronger impact on duration and occurrence than was present for onset models. Also of note, the importance of minerals only seems to appear significant mostly in cases of territorial conflicts, suggesting that location-specific nature of the resources creates an incentive for territorial secession. An alternative view would be that the presence of minerals increases the stakes for a conflict, making it last longer and become more intense in cases when the minerals could shift to a new territory. This view definitely has potential merit, and it is hard to distinguish between this explanation and the idea that minerals are useful for financing conflicts.

Oil

Onshore petroleum density (`petrodense`) has a strongly negative impact on the risk of territorial war continuation. This finding suggests that while onshore petroleum density increases the risk of territorial war onset, it reduces the risk of a long conflict. One reason for this conflict dynamic may be that a government with significant onshore petroleum deposits cannot tolerate an ongoing territorial insurgency in the region of its oil fields as continuing oil drilling may be central to that government's funding. Onshore petroleum density is significantly more likely to be associated with a war, particularly governmental war, than a low-intensity conflict. Onshore petroleum density has a strongly negative correlation with continued low-intensity conflict, which may be another indicator that while onshore petroleum can shorten local conflicts.

Increased importance of statewide oil exports (`gg_oil`) is more likely to decrease the risk of continuance of governmental hostilities. This limited finding is not as strong as expected, but may indicate the importance of oil as a dominant export for funding government operations against rebels. This finding is also limited because oil as the dominant export is a dummy variable that, when interacted with a lag conflict term, becomes collinear in the case of government war. (While this set of findings is unusual, it is consistent with the dataset as constructed, and may be partially a factor of the five-year period framework for the data.)

Proximity to Capitals and Borders and Occurrence of Conflict

For distances to capitals and borders, the most significant findings are that conflict appears most likely to continue when general low-intensity conflicts occur close to the state capital. However, low-level <u>governmental</u> conflicts will continue longer when they occur far from the state capital. I had expected that conflicts would in general last longer the further that they were from the state capital, but that effect does not appear to be present. These findings may be a product of the combination of co-linearity removing variables and changes in the spatial lag terms.

One case where the factors that lead to the onset of fighting are not necessarily the same as those that sustain fighting is the issue of distance from a state capital and low intensity government conflict. This finding suggests that when governmental low-intensity conflicts begin close to the capital – and they are more likely to do so – they will be shorter-lived because the conflict will be determined quickly. In contrast, when insurgent forces have more distance from the state capital – even if their goal is to control the government – then the conflict is likely to last longer because both sides will have more territory to cover and more challenges in reaching a decisive conclusion.

Similarly, the non-significance of borders in the duration of all types of conflicts suggests that method of modeling conflict, particularly for continuance of the conflict, may be less precise than desired. As conflicts grow and change shape, particularly if they last a significant period of time, they may expand away from borders even if it is the inclusion of border regions in their conflict space that makes the conflicts so sustainable. Thus, grid cells that are farther from borders become included in the conflict space even though borders may will be driving the conflict. Appropriately measuring and gauging the dynamic nature of conflict will continue to be challenging, although additional techniques, such as cellular automata, is a potential route.

Conclusion on Occurrence

This brief review of the results of dynamic logit approach to this dataset has revealed several instances in which the prevailing explanations for onset may not necessarily coincide with extended conflict or vice versa, such as distances from capitals and borders. In some cases, such as minerals, variables that had a less clear role in the onset of conflict have a much more significant role in the continuance of local conflicts. This quantitative approach, taking into consideration both onset and occurrence, may help to explain some of the reasons why qualitative case studies, such as those that feature minerals, can confound the proximate causes of conflict onset with the associated factors for conflict extension.

From a policy perspective, this research identifies types of conflicts that, once underway, are at greatest risk of continuing to persist unless actions are taken that would counter previous conflict patterns. For example, should a governmental low-intensity conflict start further from the state capital, the conflict might not initially appear likely to threaten the state and therefore receive less policy attention. But, based on these findings, that governmental conflict is more likely to endure than one that starts close to a state capital, which means that policymakers and advocates concerned with shortening conflict duration ought to consider focusing greater attention on such conflicts.

Moreover, this path of research opens avenues for additional future explorations of conflict spread and endurance using other techniques, such as cellular automata. These approaches, which sometimes rely on developed guidance for the paths of conflict spread, require some basis for hypothesizing the cause of conflict endurance, and this research on occurrence could be useful towards those ends as well. Ultimately, although there were a number of unexpected results and challenges using appropriate lag variables, the overall approach of disaggregate modeling conflict continuance suggests an open pathway for future research building on these lessons and expanding our understanding of conflict at the local level.

Onset and Occurrence Regressions for General Governmental and Territorial Hostilities

Variable	Base Governmental Hostilities Onset	Governmental Onset with Binary Lag	Dyanmic Logit for Governmental Hostilities		Base Territorial Hostilities Onset	Territorial Onset with Binary Lag	Dyanmic Logit for Territorial Hostilities	
			Onset	*Occurrence*			*Onset*	*Occurrence*
Spatial Lag	1.0707547***				2.3222629***			
Spatial Period Lag	-.34305749*	-0.16304439	-0.17839742	0.35859256	-1.7789383***	0.25912703	.9437051***	-0.0981221
Country size	-0.00103688	0.00094034	0.00086935	0.00212905	0.00022118	0.00068936	0.00111809	-0.00269586
Mean local rainfall	-0.00015963	-0.00020269	-0.00027886	0.00042137	-.00190384***	-0.00018785	0.00006527	0.00337601
Local forest cover vs state average	-0.16669523	-0.06792397	0.0212904	0.28592957	1.1961594**	0.01206213	.22886555*	-0.62047613
Change in local rainfall per capita	-.00782597***	-0.00536731	-.00629859**	.02086721**	.00707258*	-0.0055955	-0.00248202	-0.00651046
Local elevation range	.00136148***	.00099361***	.00148837***	-.00126536***	-0.00104577	.00114232***	.0010373***	-0.00073625
Local population (natural log)	.17825066***	.12360782***	.15761774***	-0.02246456	-.19149028***	.13657632***	.14837707***	0.1585808
Total state population	-0.000003028	-3.646e-06*	-4.300e-06**	-0.00001679	-0.00000103	-3.832e-06**	-6.301e-06***	-0.00001637
Mineral deposity density	0.02298583	0.01945196	0.02169842	-0.03921321	-.59495982**	0.02393933	0.03663272	-3.0295978**
Distance to sec. diamonds (nat. log)	-0.21871279	-0.20069445	-0.103718	0.36587331	1.3932146***	-0.1920299	-0.07018246	25.603618***
Density of onshore oil deposits	.29629452**	0.06614935	0.14162171	1.3682559	1.3362022	0.13896506	0.03721202	
Path distance to state capital	1.263E-08	-1.487E-08	-4.919E-07	-0.000000254	1.996e-06***	7.03E-09	-2.367E-07	0.000002426
Proximity to state borders	3.089e-06***	1.123E-07	8.397E-07	0.00000111	-5.155E-07	6.203E-07	1.299e-06**	2.162E-07
State military expenditure (COW)	4.328E-07	-4.629E-08	0.000001002	-0.000000158	-5.705e-06***	1.158E-07	4.319E-07	-0.00003738
State military personnel (COW)	0.00287423	0.00374123	-0.00406298	0.03639537	.02876767***	0.0029204	-0.0006566	0.22608952
GDP per capita (natural log) (GG)	-0.2602166	-.33908045**	-0.33110227	0.29122378	1.5684885	-.40259693**	-.45333551***	-0.00250785
GDP growth (GG)	-.32414994***	-.47420031***	-.5775394***	0.33707265	.34788911***	-.46716918***	-.45959118***	4.8467503
Primary commondity exports (GG)	2.3071932	0.29786877	6.3766785	-242.84478**	32.454135	0.8892884	-0.66606548	-374.44449
Primary commondity exports-squared (GG)	-7.6945697	-2.2277279	-17.751195	854.5218*	-191.86283	-3.7962007	-2.0075126	1062.5447
Oil primary export (dummy)	1.0778215	1.4934084	1.4028771	-5.6063381*	-8.0724132***	1.5701243	2.456904***	
Polity IV score (thesis modification)	.11932923**	.17834334***	.10519949**	11.368954	.19348358**	.1694046***	.09122018*	2.1264181
Density of ethnic borders (GREG)	-0.06079392	0.04121876	-0.15073151	-0.07654377	.55088329**	-0.02490769	-0.19898738	2.4232313***
Lag for occurrence of government hostilities		2.1099297***		0.29777603				
Lag for occurrence of territorial hostilities						-1.5218636		-361.50899
Constant	-3.025188	0.60682925	-1.3308569		-36.904777***	0.56438856	-1.1555108	
R-squared (adj.)	0.48754912	0.35743863	0.63955237		0.79220516	0.33396347	0.51530578	
Grid cells	83587	83587	83587		83587	83587	82379	

* p<.1; ** p<.05; *** p<.01

Onset and Occurrence Regressions for General Low-Intensity Conflict and Civil War

Variable	Base Low-Intensity Conflict Onset	Low-Intensity Onset with Binary Lag	Dyanmic Logit for Low-Intensity Conflict		Base Civil War Onset	Civil War Onset with Binary Lag	Dyanmic Logit for Civil War	
			Onset	*Occurrence*			*Onset*	*Occurrence*
Spatial Lag	1.3370114***				1.1187107***			
Spatial Period Lag	-.73583533***	-0.16265742	0.02415914	-0.21550393	-0.39523591*	0.20678572	.38362034**	-0.13372868
Country size	-0.00152067	0.00053608	0.0006755	0.00513595	0.00002814	0.0010416	0.00145259	-0.002097
Mean local rainfall	-0.00059352	-.00062724*	-.00062223*	-0.00109593	-0.00013642	-0.00014343	-0.00008444	-0.00029476
Local forest cover vs state average	-0.00507508	0.12843425	0.03629599	2.5699349***	-0.03668375	0.10171304	0.31453103	-.63286004*
Change in local rainfall per capita	-0.00460143	-0.00419141	-0.00298629	.04417402***	-0.00580974*	-0.00313467	-0.00372796	-0.0056102
Local elevation range	.00078029*	0.00024612	0.00020024	0.00046072	0.00066504	0.00055753	0.00022551	0.00078705
Local population (natural log)	0.09781522	0.07600413	0.08025735	-.24618652**	.11642827**	.09152257**	.11302465**	-0.03181252
Total state population	-7.451e-06*	-0.000007804	-0.000008316	.00001779*	-0.00000249	-0.000003311	-3.692e-06*	0.00001829
Mineral deposity density	-0.00351396	0.01811497	-0.05879979	-0.02049067	0.09083351	0.06650315	0.12202945	0.12456237
Distance to sec. diamonds (nat. log)	0.46220014	0.26575492	0.0094602	0.34444441	-0.11838157	-0.06714098	0.19234565	-0.09381881
Density of onshore oil deposits	0.56644525	0.58176081	0.52173358	1.0932465	.30206144**	0.14354846	0.16044972	-0.34490709
Path distance to state capital	4.868E-07	9.585e-07**	5.304E-07	-2.537e-06***	2.151E-07	1.635E-07	-1.389E-07	6.122E-07
Proximity to state borders	9.923E-07	-1.385e-06*	-7.018E-07	-0.000001612	3.582e-06***	8.532e-07**	1.141e-06***	4.505E-07
State military expenditure (COW)	6.361E-07	0.000000667	7.907E-07	-0.000001611	-9.230e-07*	-1.130e-06**	-1.030e-06*	3.117e-06*
State military personnel (COW)	0.01054687	0.01008988	0.01339238	-0.02228353	0.00898548	.00940436**	0.00609514	-.03589697**
GDP per capita (natural log) (GG)	0.25404522	-0.15546076	-0.12640132	1.4660582**	-0.22634763	-.35200642*	-.33772757*	-0.0830786
GDP growth (GG)	-0.04971799	-.20478416***	-.20103972***	-0.44187863	-.22718908*	-.42796441***	-.46527726***	0.22629609
Primary commondity exports (GG)	-7.2823918	9.1371941	9.3039499	-30.087612	5.6124753	7.3197673	7.1995125	-43.902865
Primary commondity exports-squared (GG)	16.280587	-22.084914	-24.223995	56.989865	-10.139776	-13.065093	-17.794939	106.01552*
Oil primary export (dummy)	-4.1149197*	-4.617853	-6.2620902	6.1407781	1.0539513	1.7499023*	1.8420968*	1.4054294
Polity IV score (thesis modification)	.15533097**	.17376508***	.12792583**	-.27386982*	0.06862461	.14745137***	0.07755159	-.24240643*
Density of ethnic borders (GREG)	-0.00569489	0.00974034	-0.07012915	0.38070014	0.01699232	0.01292856	0.06312745	-0.1166268
Lag for low-intensity conflict		-1.209762		-9.56491				
Lag for civil war					0.28922526			6.962843
Constant	-12.27822*	-5.2575132	-2.1676535		-6.0741843**	-2.5723483	-6.3767622	
R-squared (adj.)	0.45456541	0.19605024	0.34504035		0.47911609	0.30322475	0.56817478	
Grid cells	83587	83587	83587		83587	83587	83587	

* p<.1; ** p<.05; *** p<.01

Onset and Occurrence Regressions for Governmental Low-Intensity Conflict and Civil War

Variable	Base Governmental Low-Intensity Conflict Onset	Governmental Low-Intensity Onset with Binary Lag	Dynamic Logit for Governmental Low-Intensity Conflict *Onset*	*Occurrence*	Base Governmental Civil War Onset	Governmental Civil War Onset with Binary Lag	Dyanmic Logit for Governmental Civil War *Onset*	*Occurrence*
Spatial Lag	1.0575436***				.92405858***			
Spatial Period Lag	-.48732781**	0.10864542	0.13897444		-0.06275995	.3760393*	.57676771***	-0.60025019
Country size	-0.00054868	0.00112495	0.00173759		-0.00053037	0.00105499	0.00103023	0.00051
Mean local rainfall	0.00015473				-0.00022119			
Local forest cover		-0.00005456	-0.3817325	-20.990724***		0.04774355	0.13135303	2.7882892
Local forest cover vs state average	-0.16195042	-0.20371558	-0.00004533	23.737263***	-0.17066596	-0.15086302	-0.13726161	-2.9238345
Change in local rainfall per capita	-.00821322***	-.00920044***	-.00661003**	.06581579***	-0.00503629	-0.00109815	-0.00333377	0.00713064
Local elevation range	.00103824***	.00118843***	.00153209***	-.00334642**	.00142759***	.00117719***	.00102534*	-0.00075914
Local population (natural log)	.24485203***	.26655019***	.16891132**	-.5856934**	.13395752**	.08173754*	.10763632*	-0.06306312
Total state population	-9.41E-06	-.00001172*	-.00001709**	.00007239***	-3.61E-06	-5.443e-06*	-9.065e-06***	0.00002059
Mineral deposity density	-3.69E-02	-0.07079976	-.08557636**	.20428207***	.12879817*	0.1169825	.21358278**	-0.10210208
Distance to sec. diamonds (nat. log)	0.14568479	-0.06415936	-0.28415087	-0.53463943	-0.20116888	-0.1007835	0.07516865	-0.02116092
Density of onshore oil deposits	0.4307175	0.37411371	0.39086554	-1.2619801	.30838968**	.22090899***	0.1950201	-0.11662946
Path distance to state capital	-3.20E-07	1.24E-07	1.51E-07	5.271e-06***	4.86E-08	-3.00E-08	-6.53E-07	5.989E-07
Proximity to state borders	1.801e-06**	-5.97E-07	-3.15E-07	0.000001526	3.446e-06***	7.71E-07	6.72E-07	-4.193E-07
State military expenditure (COW)	1.717e-06**	2.092e-06**	2.355e-06***	-0.000001922	-1.106e-06*	-1.464e-06**	-1.307e-06*	0.0000188
State military personnel (COW)	-0.0018762	-0.00559468	-0.00470454	.16734028***	.01017774*	.01035307*	.00896472*	-0.06719615
GDP per capita (natural log) (GG)	0.25153617	0.02991427	-0.23982689	-9.8056464**	-0.38004524	-.5018191**	-.46285769**	-0.74449245
GDP growth (GG)	0.02807871	-0.10863111	-.13885412*	-0.08005379	-.45567403***	-.59758625***	-.81310952***	0.76523
Primary commondity	-5.3715903	-2.5656865	-0.32968274	-275.59051***	11.076316	11.844414	22.272349**	37.295539
Primary commodity	9.385094	2.6835686	-3.8090694	975.73618***	-25.955857	-28.137898*	-57.436144***	-196.37692
Oil primary export (dummy)	-2.8888007	-2.8066274	-3.2853447	-66.731795***	2.8685225***	3.6296042**	4.6141342***	
Polity IV score (thesis	-0.03594364	0.00432996	0.0101567	-1.5987574***	.19272891***	.22671789***	.15614958**	-.69376032***
Density of ethnic borders (GREG)	-0.20368531	-0.18735867	0.03040043	1.7347255***	-0.01503735	-0.00014711	-0.12110898	-0.02850865
Lag for low-intensity territorial		-4.804953***		64.401938**				
Lag for territorial civil war						0.53410418		4.2859542
Constant	-10.343813*	-3.5958052	1.386621		-4.4620848	-2.2787523	-5.2364582	
R-squared (adj.)	0.36504185	0.25427383	0.53532669		0.55179237	0.45961127	0.5896837	
Grid cells	83587	83587	83587		83587	83587	80671	

* p<.1; ** p<.05; *** p<.01

Onset and Occurrence Regressions for Territorial Low-Intensity Conflict and Civil War

Variable	Base Territorial Low-Intensity Conflict Onset	Territorial Low-Intensity Onset with Binary Lag	Dyanmic Logit for Territorial Low-Intensity Conflict *Onset*	*Occurrence*	Base Territorial Civil War Onset	Territorial Civil War Onset with Binary Lag	Dyanmic Logit for Territorial Civil War *Onset*	*Occurrence*
Spatial Lag	2.2728658***			All of the occurrence results for Territorial Low-Intensity Conflict were removed do to collinearity	3.1210658***			
Spatial Period Lag	-1.3392701***	-1.2000907***	-0.44212311		-4.5161123***	-5.1304408***	-0.26630046	0.82401056
Country size	-0.0031916	-0.00131788	-0.0001976		.00451322***	.00653973***	.00368822***	-.09320803**
Mean local rainfall	-.00249538***	-.00313564***	-.00264196***		-.00185288**	-0.00036255	-0.0004454	-0.00053605
Local forest cover vs state average	1.0404073**	1.3263448***	1.1927739***		0.10658342	1.9244627***	1.5192259**	-3.796574***
Change in local rainfall per capita	0.00515139	.00941103*	.0108758*		0.00789443	-0.00907771	-.01374958*	-.02006844**
Local elevation range	-0.00026286	-0.00044567	-0.00120182		-0.00187296	-.00528945***	-0.00224466	.00692918***
Local population (natural log)	-.20020976**	-.24922674***	0.01322596		-.11648904**	.18479334***	.30419129***	-.17633267*
Total state population	-0.000006382	0.000002988	0.00000548		.00003885***	0.00000642	0.000007474	-.00039225**
Mineral deposity density	-0.46492641	-0.57670649	-.65083781**		-0.52858795	-1.5030397***	-.77418613*	2.4263044**
Distance to sec. diamonds (nat. log)	1.9767617***	1.1417965	1.3706552***		1.5480438***	1.7781994**	3.0872277***	-3.1829604***
Density of onshore oil deposits	0.68324179	1.0551758	0.73538148		0.9364764	4.6336051***	3.5119341***	-5.0128187*
Path distance to state capital	1.088e-06*	1.057e-06*	3.777E-07		2.381e-06***	4.225e-06***	2.941e-06***	-1.941e-06**
Proximity to state borders	-0.000001638	-4.026e-06***	-1.958e-06*		9.461e-06**	0.000002343	0.000001192	3.644E-07
State military expenditure (COW)	-.00001312***	-8.271e-06**	-2.342e-06***		-6.454e-06*	-0.000001327	0.000001292	.00022378**
State military personnel (COW)	.06171472***	.0477661**	.02710921***		-0.00679519	0.03196838	-0.00840627	-.71965535**
GDP per capita (natural log) (GG)	.9450147***	-0.30149518	0.0391822		2.570675	0.80063149	-.30483568***	-120.05848**
GDP growth (GG)	-.29460843**	-.40764202**	-.25104106**		1.6035876***	1.12058***	.58564634***	-18.474554**
Primary commodity exports (GG)	-2.2128812	85.149352**	89.769845**		11.966785	65.145656	80.621417	139.58957
Primary commodity exports-squared (GG)	-45.097168	-261.06494*	-280.02843**		-45.906792	-189.79503	-386.27183	930.7716*
Oil primary export (dummy)	-0.70879176	-5.0234334*	-8.3540906***				-29.892063	328.68078
Polity IV score (thesis modification)	.6704225***	.53427834***	.42323896***		-.83428372***	-.46916296**	-.40461421*	-1.6513138*
Density of ethnic borders (GREG)	.92948061***	.48426294**	-0.17146362		-.3481611*	.28833175**	0.22936759	0.44286903
Lag for low-intensity territorial conflict		2.2557371						
Lag for territorial civil war						3.5910512		991.79077**
Constant	-36.519412***	-18.175885	-25.670057***		-73.216776***	-57.451437***	-59.903661***	
R-squared (adj.)	.79939735	.63112563	.52751431		.90630008	.72685961	.82453372	
Grid cells	83587	83587	82541		63877	63877	83587	

* p<.1; ** p<.05; *** p<.01

CHAPTER 6: CONCLUSION

For much of the past decade, there has been a concerted effort to explain "civil war" through empirical analysis of political, economic, and geographic data. In some cases, the results have been compelling – such as the finding that civil war at a state level is closely related to previous civil wars, low GDP per capita, slow economic growth, and a large population (Hegre and Sambanis 2006; Blattman and Miguel 2009). Still, several key issues have remained largely unresolved, such as the role of extractive resources, and even for those issues where the results are recognized as robust, the findings do not necessarily lead to conflict resolution as much as conflict prevention (Hoeffler 2009). More importantly, research to date has rarely ventured to consider low-intensity conflicts with fewer than 1000 battle-related deaths per year and rarely differentiated conflicts based on the intents of the conflict actors, whether to control territory or the government. For this project, the data was tested with a range of measurement approaches (see Annex II), and a analytical framework was used that differentiated insurgencies that seek to control a set of territory and those that seek to control the government, as well as differentiating civil wars from lower intensity conflicts.

My approach differs significantly from the strategies employed by major studies over the past decade. The Collier-Hoeffler framework for civil war, one of the most often cited frameworks for studying civil wars, suggests that greed and grievance are both components of civil war, and uses a common regression model to explain a wide range of civil wars, although it does not differentiate any sub-type of conflict, such as "ethnic" war or wars of secession versus wars to control the government (Collier and Hoeffler 2004).

Fearon and Laitin do differentiate between types of conflict in their study, noting differences between civil war in general, "ethnic" wars, and civil wars that include rebellions against empires (Fearon and Laitin 2003b). Buhaug and Rød use the Uppsala conflict dataset as their basis for conflicts, thereby differentiating governmental and civil wars largely on the basis of geography, but they analyze all civil wars, leaving out lower intensity conflicts and using such a large timeframe – over three decades – that political and economic variables are largely excluded from their analysis (Buhaug and Rød 2006). The analysis conducted for this project recognizes that there is not one single explanation for civil wars, or all civil conflicts. Rather, at the local level, different types of war have different causes, and lower intensity conflicts have another similar but separate set of correlated factors.

Moreover, this project has sought to use different measurement approaches and interaction terms to better differentiate some of the theoretical explanations of conflict, since linking theoretical explanations and empirical measurements has been an ongoing challenge for the field (Hoeffler 2009). In cases such as the role of borders, where multiple theoretical explanations exists for the correlation between conflict and proximity to international borders, this project tested interaction variables across a number of potential explanations. Proceeding with the borders example, evidence emerged that one of the most significant relationships are when borders have more passable terrain, suggesting that easier movement of people and goods across borders is central to their significance, rather than having mountains that might provide havens for rebels. Similarly, in an effort to connect variable specification with theoretical explanations, the project tested a mid-range set of elevations for the likelihood of conflict onset, and the significant results for conflict in general suggests that conflict is most likely to onset in areas that do not have extreme elevations, as some more linear specifications of the variable have suggested. On a related note, elevation and straight-line distance were incorporated into measures for distance from state capitals and borders to provide a more realistic sense of the distance between a grid cell and the relevant feature. These efforts to use technology improve the specification of variables and better connect them with theoretical explanations for conflict onset and occurrence.

The results from this project help to differentiate conflicts by their intent and intensity, providing some explanations for where disputes in the literature may be emerging. For example, while Collier and Hoeffler assess both greed and grievance in a common model, these differentiated results suggest that greed and grievance (or opportunity and motivation, as other models describe them) correlate with different types of conflict. The traditional elements of grievance – lower GDP growth with lower levels of local vegetation (as a proxy for growth) – are more often found in cases of governmental conflict. The traditional elements of greed – stronger GDP growth and more vegetation at the local level – are more often correlated with territorial conflicts. These divisions are not absolute; the local presence of oil in governmental conflicts suggests greed elements and the lower levels of local rainfall in territorial conflicts suggest possible grievance elements. Nonetheless, the core argument holds: different types of conflict have significantly different correlates that require explanations and cannot being easily unified into a single group for all purposes.

Also, while the existing literature has found robust results on civil war, conflict prevention – a central focus of this field of research – may be significantly strengthened by identifying the causes of low-intensity conflicts to ensure that they do not metastasize into larger wars. In focusing on these lower intensity conflicts, it becomes clear that previous aggregate analytical approaches, using a single conflict type and statewide variables, are not always as useful because there appear to be no common significant non-lag variables in the onset model for low-intensity conflict and high-intensity civil war. The finding that local variables, such as local elevation range for low-intensity conflict in general or ethnic diversity for territorial low-intensity conflict, are relevant suggests that a disaggregate approach to studying low-intensity conflict may have value. This finding suggests that while the existing literature may have a reasonable aggregate approach for studying wars, a more localized, disaggregate approach is helpful for studying low-intensity conflict.

For both the academic and the policy-making communities, these findings suggest that a differentiated approach to conflict prevention is appropriate. Academic studies and development programs that seek to explain or ameliorate all forms of conflict in a single approach are unlikely to address the fundamental distinctions that exist between conflict types. While the country-year approach has had value for economic studies, particularly those that discuss the likelihood of governmental civil war, it is not a sufficient approach for explaining a much broader set of conflicts, some of which – such as territorial low-intensity conflicts – are ill-suited to the aggregate approach. In the policy community, this overall finding suggests that government aid programs that broadly target "conflict prevention" may need to become more targeted to specific conflict dynamics. In the specific case of defense, this compiled evidence of significant distinctions between classes of conflicts, particularly between different intensity levels, suggests that development of combat programs that target fighting broadly, rather than which are attuned to the specific context of a low-intensity conflict or more high intensity civil war will not be sufficient. While defense communities already recognize civil conflicts as distinct from past interstate conflicts, and it is not adequate to paint all civil wars with the same brush.

Using statistical analysis to break down a process as complicated and messy as violent conflict or war will never be easy, but differentiating types of hostilities and including more detailed, localized information can improve the approaches that researchers use to track conflict trends. Efforts at intervention and conflict prevention can be better tailored and targeted with a more specific understanding of the nature of the causes of the type of conflict and the role that local factors play in the onset and continuance of fighting. For conflict prevention efforts to be strengthened and focused, research that recognizes differentiation of conflicts and local factors is necessary.

ANNEX 1. REPLICATION OF BUHAUG-RØD MODEL OF CONFLICT IN AFRICA USING GWR

This replication of one of the Buhaug-Rød models for conflict in Africa will serve both to show the effects of Geographically Weighted Regression (GWR) on results but also to highlight the detailed coding issues involved in their dataset. The guidance and knowledge gained from this replication and the process of reviewing their data has informed both the method by which data was derived and the variable selection for the models that constitute the main study. Before going into the details of the replication, it is useful to provide some context for the endeavor to identify local statistical relationships with GWR.

Determining Local Statistical Relationships

While the first two changes from the process followed by Buhaug and Rød focus on changing the dataset by adding variables or altering their representation, the last change that this procedure proposes is to perform a different kind of analysis. Buhaug and Rød perform a traditional classic regression on the spatially disaggregated data that they have collected. While there is nothing that is analytically incorrect with this approach, it may not represent a number of interesting trends in the data that would be captured by generating local statistics through a process called Geographically Weighted Regression (GWR). GWR analyzes data incorporating its location into the regression, thereby generating local statistics – coefficients – for the various independent variables included in the regression (Fotheringham et al. 2002). GWR performs regressions at each of the locations over a given geo-referenced dataset and weights the importance of data points based on its distance from the regression point. The result is a map of different statistics rather than a single statistic for an entire region or the world. Visually, the distinction can be understood through the two maps below which reflect the difference in the detail of the analysis between generating a single statistic for a whole are and generating a specific statistic for each location in a dataset (Fotheringham et al. 2002)

Data points are assigned weights according to a gradient, ranging from 1 down to zero, and the gradient can be changed by assigned different "bandwidths." A high bandwidth will create little difference between the weights assigned to points close to the regression point and those further away, while a low bandwidth will create a sharper decline from the weights of near points to far points. There is no ideal bandwidth setting – in some cases a lower bandwidth may be appropriate, in some cases higher bandwidths may produce better estimates. The process of optimizing the bandwidth to reduce the standard errors of the estimates occurs for each geographically weighted regression conducted. In some preliminary testing of data and methodology, it is clear that bandwidth determination will be a crucial factor in the results. As can be seen in the results presented below, changes in bandwidth can affect both the sign and significance of results in different locations (Haring-Smith 2008).

In an example unrelated to conflict offered by Fotheringham et al. to explain the need for local statistics is for house prices in England: older houses are more valuable in rural areas where they add charm and history, whereas older houses are less valuable in urban areas because they lack modern amenities and require more repair. Looking at them overall, one might conclude that the age of a house has no effect on its price based on a classic multivariate regression because the two effects offset each other. With a geographically weighted regression, local statistics can be generated that show that there is a positive coefficient for the age of a house in rural areas but a negative coefficient for a house in urban areas by performing many spatially referenced regressions using only the portion surrounding data from the entire geo-coded dataset. The same concept for housing prices is also possible with conflicts. Many aggregate regressions have not provided a clear or robust relationship between non-fuel mineral resources and civil war (Ross 2004b; Hegre and Sambanis 2006; Ross 2002; Humphreys 2005), or they fail to address non-fuel minerals at all (Collier and Hoeffler 2004). This mixed picture of findings runs counter to much theory that mineral resources provide a source of funding for rebels, and even some case studies that have documented the use of diamonds to fund rebel operations (Smillie et al. 2000; Le Billon 2001a; Malaquias 2001). Some disaggregate studies and case studies have found a relationship between lootable forms of diamonds and civil war (Lujala et al. 2005; Snyder and Bhavnani 2005; Lujala 2010). From developed theories of conflict, though, there may be a local statistical explanation of the findings. It could be that when mineral resources are located close to government capitals, they act as deterrents on civil conflict because they provide a source of funding for government military forces, but when located far away from government capitals, they provide a reasonably controllable source of funding for rebels. Thus, while a classic regression might conclude that there is no relationship between diamonds and conflict, a geographically weighted regression might conclude that there are in fact two relationships – one positive and one negative – but that they offset each other in classic regression analysis.

The value of local statistics is not just at the level of African data – it can also apply if the study is later expanded to include regions other than Africa. For instance, in disaggregated analysis of conflicts, we have generally found that conflict polygons in Africa have been associated with lower levels of rough terrain – both in terms of mountains and forests – but the finding is limited to Africa (Buhaug and Rød 2006). Globally, we have fairly robust evidence that rough terrain is associated with conflict (Fearon and Laitin 2003b; Hegre and Sambanis 2006). It is entirely possible that both of these findings are accurate – that conflict in Africa tends to occur in less mountainous areas because rebels lack the road system or reliable food supplies in mountainous or densely forested areas to make it possible to sustain a significant rebellion, but, when the globe is taken as a whole, the improved road quality and food distribution networks make it possible to sustain rebellions in areas with rough terrain.

Given that geographically weighted regressions have the potential to present different relationships in different places, it could increase the threshold of theorizing needed *a priori* to make testable hypotheses. Even with extensive development of hypotheses, it is still likely that geographically weighted regression will produce results that predicted conflict in some areas but not in others. Part of the value of geographically weighted regression is that it identifies locations with particularly strong findings or anomalies, allowing researchers to then use this quantitative method to narrow the focus of their qualitative research.

It is also a central element of some of the analysis of conflict in Africa that there exists a distinction between territorial and governmental conflicts. Territorial conflicts tend to be located further from state capitals, in more remote areas, closer to borders, with less dense populations, while governmental conflicts are further from borders, closer to capitals, and in areas with more dense populations (Buhaug and Rød 2006). Using local statistics in might be possible to generate some form of general model for conflict in which the distinction between different types of conflict is removed.

An extension of the basic concepts of GWR is mixed GWR – which mixes variables that have non-stationary influence and those that have stationary influence. For instance, in the example of housing prices, a variable with non-stationary influence might be garages, while a variable with stationary influence might be the unemployment rate in the surrounding area. Having a garage adds significant value to a house outside London, but close to central London, a garage adds little value. In contrast, higher unemployment rate would have a constant negative influence on property rates regardless of the area (Fotheringham et al. 2002). In the area of conflict studies, it is likely that diamond deposits should be designated as a non-stationary factor, but while democracy scores might be better designated as a stationary factor. In theory, the location of a diamond deposit affects its ability to spur conflict by either funding the government military or the rebels, but democracy could theoretically have the same effect on the likelihood of conflict regardless of location. On the other hand, perhaps democracy inhibits conflict in locations far from capitals but encourages it in locations nearer to capitals. It is not immediately clear whether a variable such as democracy should be considered as a stationary or non-stationary factor, and thus different models using democracy in each form will need to be compared. Using the technique of mixed GWR, it should be possible to integrate variables whose influence will not vary much over the territory of a state, such as democracy or economic growth, with greater consistency between the theories provided and the statistical methods used.

While GWR should be useful for generating local parameter estimates for variables, it should also be useful for determining local standard error statistics. The reality is that there are some geographical areas where data collection is so sparse that reaching reliable parameter estimates is not possible. With a classic regression, the same level of certainty is presented for the entire dataset, but with geographically varying data, it may well be that there are some areas that are significantly more certain and have smaller standard errors than other areas. This information is important in providing both clarity and policy advice from a study's findings. It is easy to intuitively understand that some areas – such as central Russia – where limited data exists may result in greater uncertainty about the parameter estimates, but a classic regression model does not present any indication of the degree of greater uncertainty – the standard errors for each parameter apply across the entire space of the model. By generating geographically weighted regressions centered at points throughout the space of the dataset, it is possible to create a measure of the level of uncertainty for each parameter estimate in each location.

Applying GWR to the Buhaug-Rød Model of Conflict

Using Geographically Weighted Regression changes the nature of the analysis being performed. Rather than assuming that a variable has a uniform impact across the entire space of a dataset, it allows the variable to have different impacts in different places. While the overall aim of this study is to show how GWR and combining aggregate and disaggregate data could lead to significant results, it is important to demonstrate how the methodological change alone impacts results. To demonstrate the effect of GWR, this study will briefly focus only on the Buhaug-Rød dataset that was used for their "Local Determinants of African Wars" article in 2006 (Buhaug and Rød 2006). This dataset, which comprises 3206 grid cells of 100km x 100km covering the continent of Africa tracked a number of variables – population density, proximity to international borders and to state capitals, proximity to diamond and petroleum deposits, mountainous terrain, forest, whether language in a location differed from language in the capital, and road density.[19] By using GWR instead of ordinary least squares regression, this chapter will explain the differences that we can see emerging from determining local statistics.

Buhaug and Rød run eight different regressions – four models (Relative Location, Remoteness, Resources, and Identity) run over two different independent variables (Government Conflict and Territorial Conflict). While it would be possible to run this replication with any of the Buhaug-Rød models, this replication uses the Resources model for Territorial Conflict. The reason for selecting this model is that its findings when run by Buhaug-Rød were not particularly conclusive (see highlighted Model 3) in the table below. The distance to petroleum was not significant, and the distance to an international border and the distance to diamonds were only weakly significant. This may be partially explainable because increased distance to the capital had such a strong and significant influence. Still, given the substantial aggregate analysis that suggests a relationship between conflict and resources, the non-findings are surprising. It warrants investigation regarding whether there maybe offsetting local factors or if the role of resources remains unclear from the data.

[19] The number of grid cells used in any approximations, though, could be fewer than 3206. For instance, the territorial conflict models run below use 2885 cells.

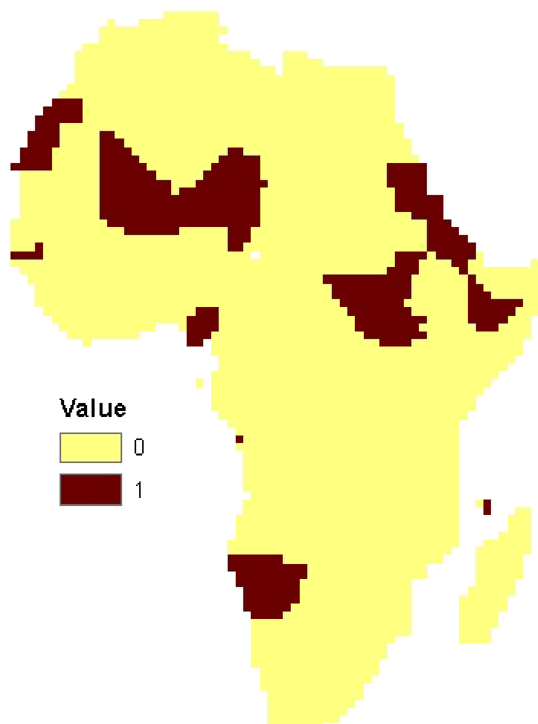
Logit estimates of onset of territorial conflict

	(1)	(2)	(3)	(4)
Spatial lag	3.692 (3.77)***	4.199 (4.66)***	3.427 (3.28)***	3.660 (3.88)***
Distance to border[a]	−0.182 (2.09)**	−0.252 (3.35)***	−0.102 (0.88)	−0.157 (1.83)*
Distance to capital[a]	1.360 (3.13)***	1.015 (2.00)**	1.378 (3.17)***	1.321 (2.94)***
Population density[a]		−0.066 (0.67)		
Relative road density[a]		−0.404 (1.89)*		
Mountain[a]		−0.276 (2.19)**		
Forest[a]		−0.267 (1.24)		
Distance to petroleum[a]			0.215 (0.53)	
Distance to diamonds[a]			0.941 (1.68)*	
Minority language				0.676 (1.33)
Intercept	−10.692 (3.86)***	−7.091 (2.06)**	−18.785 (3.18)***	−10.932 (3.75)***
N	2885	2885	2885	2885
Log pseudolikelihood	−922.1	−847.6	−842.0	−906.9

Note: Regression estimates with robust *z* scores in parenthesis (standard errors clustered on countries). $*p < 0.10$; $**p < 0.05$; $***p < 0.01$.
[a] Logged.

The Data on Territorial Conflict

Observed Values of Territorial Conflict for Resource Model based on GWR of Buhaug-Rod dataset (3.4 bandwidth)



The grid to the right shows the values for territorial conflict used in the regression from 1970-2001. The data here represent territorial conflict onset, and they range over a number of conflicts from the battle for Western Sahara between Morocco and the Polisario Front to the fights between Comoros and the Republic of Anjouan for control of the islands. The gridding here follows the data provided and used by Buhaug and Rød, although there are some notable absences. For instance, there are no conflicts marked for the entirety of the Democratic Republic of the Congo. This is partially because several conflicts (Zaire against the FLNC from 1977-78; the DRC against the AFDL from 1996-97; the DRC against the RCD and MLC from 1997-2001) are marked as governmental conflicts rather than territorial conflicts. While that classification may be debatable, it is driven by the stated aims of each insurgency, which all sought to capture the government in Kinshasa. Also, contrary to what this thesis will aim to do, conflict is merely marked as present or absent, with no consideration of a conflict's time of onset or overall intensity.

Overall Fit of the Model

When the data provided by Buhaug and Rød is run through a traditional OLS regression, the r-squared fit is 0.23. (For their work, Buhaug and Rød run the variables in a regression that clusters errors by country, so their results are slightly different – measured by a log pseudolikelihood of -842.0.) Because the GWR model fits data more closely to the results by giving more weight to data near the target point, the GWR model has an r-squared fit overall of 0.83. As can be seen from the map of predicted values from the GWR (at the end of the chapter) and the map of predicted values from the logit regression by Buhaug and Rød (below), the errors are much smaller, particularly throughout central Africa. This fit is markedly better than the traditional OLS regression.

At the end of the chapter are two maps that show the predicted likelihood of territorial conflict and the residuals from the actual predictions from the GWR. The prediction is made as a linear function, so there are some negative predictions and some predictions higher than 1. Because the GWR process draws so heavily from the data around a local area, it is notable that all of its predictions for central Africa are very similar because there is minimal variation in these areas. In almost all of the grid cells in central Africa, there is no recorded territorial conflict, so GWR process is drawing on static values for all of its most heavily weighted cells. (There is some variation in the predictions, but this is largely based on the conflict cells located further away, which have considerably less weight in the prediction.)

The residuals map for the GWR are very cleanly lay out the differences between the predicted and observed values, and it is noteworthy that they almost outline the observed values – with high positive residuals in areas where conflict occurred and high negative residuals in areas next to where conflict occurred. This result is to be expected because the GWR is indicating that an overall area is at risk of conflict, but the conflict is defined in a binary manner – either present (1) or absent (0) – so the residuals naturally take opposite signs on either side of the conflict boundaries.

It is notable that there are some areas where the predicted value of conflict is so high that there are very small (or even slightly negative) residuals. As will be seen later, this result is partially because the local factors in eastern Mali – distance from the capital, proximity to international borders, and an absence of diamonds. It is also notable that the distance to a border or the capital can make some zones of a territorial conflict more conflict-prone than others. This is reflected in the smaller residuals on the parts of conflicts that are further from capitals – for instance the southern portion of Western Sahara or the northern portion of Niger. These areas also correspond, roughly, with the bases of power for the insurgent groups – which are naturally basing themselves in areas of strongest support and least ability of the government to reach them. These findings are all consistent with the broad findings of Buhaug and Rød and the aggregate findings of Fearon and Laitin (Fearon 2003; Buhaug and Rød 2006).

It is also interesting that in one of the largest concentrations of conflict – in northeast Africa – the residuals are consistently high. While the residuals were marked in all of the areas of conflict, it is notable that they are some of the highest and almost perfectly align with the conflicts outlined, with no smaller sub-areas that are more accurately explained – as in eastern Mali. This could be an indication that while the model has some explanatory power for the conflicts in Mali and Niger, it is less helpful in explaining the conflicts in Sudan and Ethiopia. One of the apparent reasons for the divergent explanations of conflict is that Buhaug and Rød coded a substantial spatial lag in eastern Africa, where spatial lag was based on the amount of conflict covering the territory of a country, and lower spatial lag values in Mali and Niger. While it is not entirely clear why the spatial lag values are coded as low as they are in Mali and Niger, it affects the two different predictions because spatial lag (the variable `terrlag`) has strong coefficients with significant t-values.

Overall Findings

The overall findings of the GWR model are similar to the Buhaug-Rød regression. As with Buhaug-Rød, distance to a capital has a markedly positive relationship with conflict with distance to diamonds have a weaker positive relationship overall, while distance to petroleum and distance to an international border have more unclear overall results. The table below summarizes the 3206 parameter estimates for each variable that the GWR process creates. While it does not offer the detail of the maps that are included in the annex, it provides an overall sense of the findings of the GWR process, since seeing general conclusions from the maps can be difficult.

Table of variable names

```
lnborddi     natural log of the distance from the grid cell to an
international border
lncapdis     natural log of the distance from the grid cell to a state
capital
ldiadist     natural log of the distance from the grid cell to a diamond
deposit
lpetdist     natural log of the distance from the grid cell to a petroleum
deposit
terrlag      territorial lag variable (by Rød-Buhaug method)


*********************************************************
*          PARAMETER 5-NUMBER SUMMARIES                 *
*********************************************************
   Label       Minimum  Lwr Quartile        Median  Upr Quartile
Maximum
-------- ------------- ------------- ------------- ------------- ------------
-
```

Intrcept	-9.138518	-0.951862	-0.077167	-0.000292	6.764615
lnborddi	-0.196813	-0.008598	0.000063	0.006621	0.108800
lncapdis	-0.286913	0.000287	0.020156	0.121539	0.688735
ldiadist	-0.817339	0.000009	0.004022	0.033330	0.462720
lpetdist	-0.610843	-0.038024	0.000021	0.013345	1.252045
terrlag	-0.709321	0.134866	0.812983	2.005773	34.193306

From this table of results, it is clear that spatial lag (`terrlag`) is the dominant variable, with a generally consistent relationship with conflict and high coefficient values, and so the difficulties with its coding become more critical to the analysis of each local result. Spatial lag is discussed in detail below because of its significance, and some issues with its coding are raised.

Spatial Lag

The coding of the spatial lag variable is described by Buhaug and Rød as the following:

> To account for spatial autocorrelation, we generated a country-level spatially lagged dependent variable…which measures the number of years with conflict in any other part of the country, divided by the total number of years of observations for the cell. The lag thus takes on values

> between ‘0’, indicating no conflict in other cells in the country in the sample period, and ‘1’, which implies that some other areas of the country hosted a conflict in every year between 1970 and 2001. (Buhaug and Rød 2006)

The reason for so fully quoting their description of this variable is that its coding is critical to the outcomes of the model both in their 2006 analysis and in this replication using GWR. Because Mali hosted relatively few conflict-years (only 1990 and 1994 according to their data), it has a very low spatial lag value – less than 0.1 – but all of northern and eastern Mali experiences a conflict in the period under study – so the entire region receives a conflict score of 1. Meanwhile in Ethiopia, which experienced conflict-years from 1975 until the end of the period of study, the spatial lag value is over 0.8 – but the conflict regions still get a conflict score of 1. As a result, the territorial lag variable has a very high coefficient in the area around Mali while it has a much lower coefficient in the area around Ethiopia. (See the maps at the conclusion of the chapter for a demonstration). Another result is that the other variables – such as diamond and petroleum distances – have much more significance in the areas around Mali and Niger than around the Horn of Africa.

What also emerges from this analysis is that the method employed by Buhaug and Rød to measure conflicts is sub-optimal at best. Although conflict is coded at a disaggregate level, the territorial lag variable was generated at the aggregate level, creating distortions in those areas of countries that experienced little conflict. This creates significant distortions that gives territorial lag divergent impacts in different areas when there is almost no theoretical or empirical justification for this finding, which is essentially a product of Buhaug and Rød's data specification on the variable rather than an underlying theoretical reason. A variable to measure the true distance or the density of conflict cells or battle deaths in a given area would be more appropriate.

Distance to an International Border and Distance to the Capital

Distance to an international border has a negative but not significant relationship with the likelihood of conflict in the Buhaug-Rød model. Based on existing literature, it would be expected that the closer a location is to an international border, the more likely conflict would be because of the easy availability of refuge across the border, the possibility of supportive members of a common ethnic groups in the neighboring state, or the challenges of projecting state influence across borders. (While it might initially seem that the distance to an international border and the distance to the capital might be mostly an inverse of each other, the different placements of capitals in countries means the two variables have a correlation of $r = 0.0028$.)

The map of the t-values from GWR of the effect of distance to an international border on the likelihood of conflict reflects that there are areas where the distance to the border has a positive relationship and areas where it has a negative relationship. What is more concerning, and actually reveals some of the weakness of the variable itself, is that the map of its relationship with conflict does not track cleanly against the borders of states. In other words, there are areas where any value for the distance to an international border increases the likelihood of conflict, regardless of the distance to the border. Unfortunately, because of the heavy weight given to local data in this method, this more likely indicates simply that an area is conflict prone rather than indicating that the distance to an international border has a compelling relationship to conflict.

Also consistent with the Buhaug-Rød model, the distance to the capital has a positive relationship with the likelihood of conflict. There are very few areas of the GWR results map that reflect a negative association, but it is noteworthy that the coefficient does show some variation in its strength. In other words, while the coefficient for the distance to the capital in consistently positive across the Sahara and the Horn of Africa, it becomes markedly stronger in areas where conflict is observed. This variation suggests that the distance from the capital was a significant and strong contributing factor to the conflicts observed in these areas. This variation in the strength of the coefficient also corresponds closely with the actual distance from the state capital – with stronger coefficients in areas that are markedly further from capitals and still positive but much weaker coefficients in areas close to capitals – particularly along southern Morocco and in Mali and Niger. This finding bolsters the existing literature from aggregate and disaggregates studies of territorial conflicts that suggests the state's capacity to project power is limited the further than a territory is from the state capital.

There are two noteworthy instances of strong negative coefficients for the distance from the capital. The first is the area in southwestern Algeria, far from the capital Algiers, where the distance from the capital is significant, but it also has a marked negative effect on the likelihood of conflict. From case research and existing research, it seems much more likely that the reason why territorial conflict is not more likely there is because of the secondary diamond deposit in the region, not its proximity to the capital. The second location is the area outside of Angola's capital – in this case again, an area somewhat far from the capital is marked has having a negative impact on conflict. In this case, similar to Algeria, a strong resource effect is also likely an explanation. The distance from petroleum has a strong positive relationship with the likelihood of conflict in Angola and Namibia, and in this case, the distance variable appears to be cited for the lack of territorial conflict in Angola, while it does exist in Namibia.

Overall, this replication of the Buhaug-Rød data with GWR confirms the earlier findings – that the distance to an international border has an ambiguous effect on conflict but that the distance to a capital is a strong determinant of conflict. The replication does introduce some wrinkles though, as the mapped results make it much easier to see local interactions between relative location and mineral resources.

Distance to Diamonds

The result for distance to diamonds shows a positive and weakly significant relationship in the Buhaug-Rød model, suggesting that the further away from diamonds the more likely of being prone to territorial conflicts. This result may be because the government invests more resources in securing its interests around diamond deposits or because diamond deposits lead people in those regions towards governmental conflict, rather than secessionist territorial conflict, as Buhaug and Rød suggest (Buhaug and Rød 2006).

In reviewing the results from the GWR, it is not as clear that diamonds had a significant positive relationship with deterring conflict. While the median coefficient for distance to diamonds is positive, the coefficient is very weak for much of the continent. The only two areas where the coefficient shows marked strength is in northwestern Mauritania and a small portion of northern Mali, where there is a much more negative coefficient, and a significant portion of Ethiopia, where there is a much more positive coefficient. Since neither of these areas have any diamonds, it is not immediately clear what is leading the two regions to have such divergent coefficients. The best likelihood appears to be related to the sizeable distance to any capital from northeastern Mauritania and northern Mali and the relative proximity to several capitals in the smaller countries around Ethiopia. In other words, the regression results may be suggesting that diamond deposits located further from state capitals have a more deterrent effect on civil war by encouraging the state to project resources in that direction, while diamond deposits near state capitals have the possibility of increasing conflict as it increases the value of the "prize" of capturing the state. While this regression tracks territorial and not governmental conflict, this explanation of the results would be consistent with the findings that "diamond-abundant areas have a significantly higher probability of hosting a governmental war than less affluent parts of the country" (Buhaug and Rød 2006).

Distance to Petroleum

While diamonds had a negative relationship with conflict in northwest Africa and a positive relationship near the Horn of Africa, petroleum appears to have the opposite effect. In the pocket of the western Saharan region that does not have oil – western Mauritania and Mali in particular – there is a marked positive relationship between the distance from petroleum and the likelihood of conflict. In the area near and around the Horn of Africa, where there are more oil deposits, the distance to petroleum appears to have a positive impact on the likelihood of conflict. In these situations, it may be that the lack of any major state resources in the outskirts of a country makes it more prone to conflict because of the absence of state presence and investment there, while the presence of resources closer to a capital creates more incentives for groups to seek to capture portions of the state.

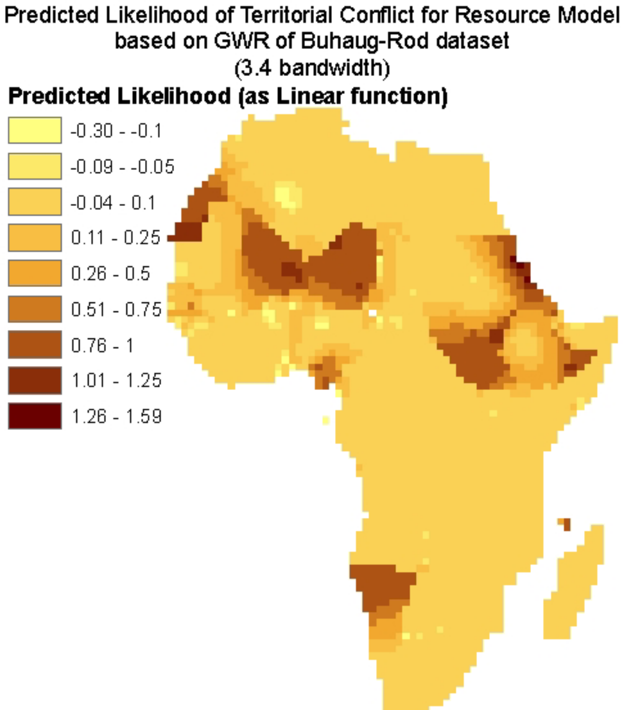
If this underlying picture were true – that resources on the outskirts of a state actually reduced the chances for conflict by increasing state investment and presence in the region while resources closer to capitals increased the chances of conflict by making the state more valuable as a “prize” and potentially reducing the need to reach out to outlying areas to fulfill the state’s resource needs – it would be a marked shift from the model that suggests that outlying resources create funding for secessionist groups on the edges of a state.

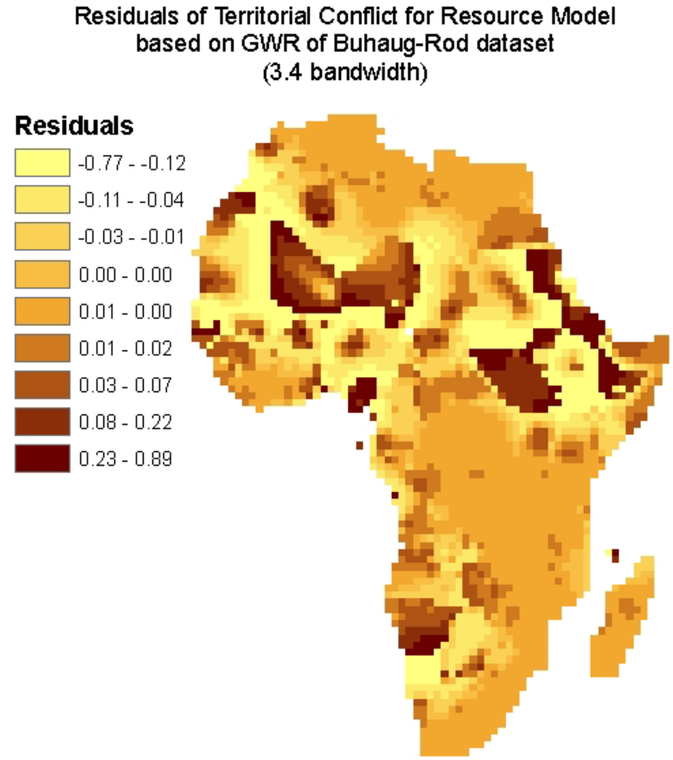
Conclusion

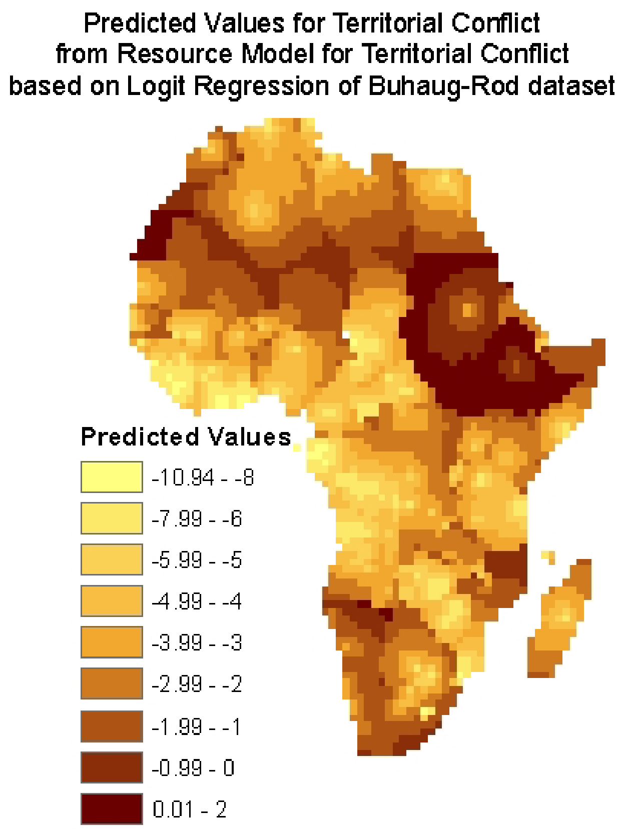
From this replication of the resources model for territorial conflict and the detailed examination of the variables that were involved, we can draw several key points: 1) the state-level definition for spatial lag had strong explanatory power, 2) the effect of remotely-located resources appears to have generally deterred territorial conflict, and closely located resources may have spurred territorial conflict but broader examination is warranted, and 3) because GWR reduces residuals so effectively by heavily weighting local data, it is difficult to use GWR as a predictor for conflict in areas not already experiencing it.

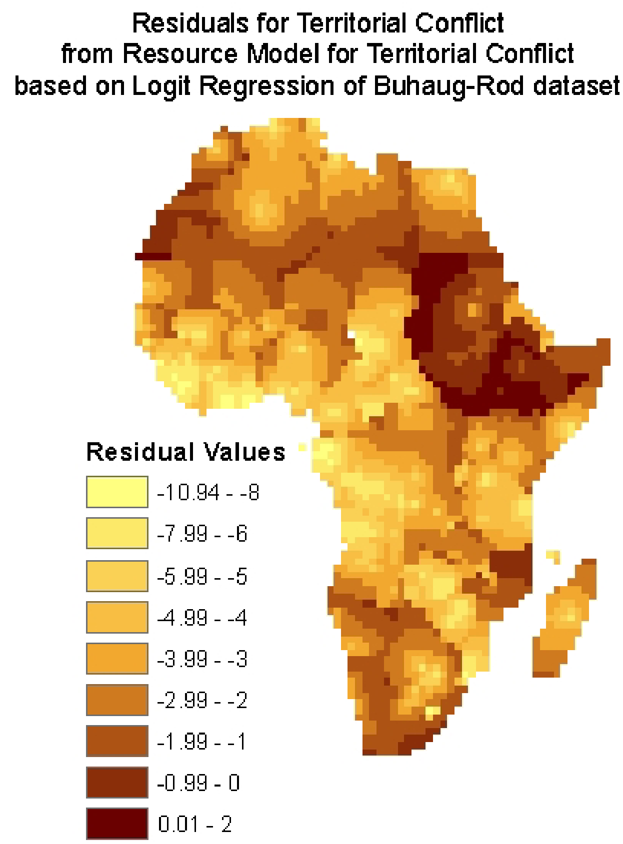
One of the downsides of GWR is that it does not do as good a job as one might hope for identifying unexpected locations for possible future conflict. Because the approach draws heavily on local data, it tends to accurately predict areas of consistent non-conflict. As a result, while the GWR approach may illuminate some of the reasons why a conflict occurred (or did not occur), it is not as effective in identifying areas likely to experience future conflict but currently at peace. The underlying theoretical issue that creates this difficulty is that GWR assumes that a variable has different impacts in different places. So, when diamonds are absent from southern Sudan it is marked as a significant contributor to conflict, but when they are absent in northeastern Sudan, it is marked as a significant deterrent to conflict. While this shift is interesting when considering what may have contributed to the different conflicts in Sudan, it is not as helpful when the regression indicates that all of the diamonds in central Africa deterred territorial conflict, but there is also no territorial conflict marked in central Africa. In the process, smaller anomalies are explained because the local data around them is so significantly weighted – but when searching for small anomalies such as areas currently at peace but "shouldn't be" based on their basic conditions, it is more helpful to have a global approach to the data. Thus, the global regression determines that generally diamonds have a given effect on the likelihood of armed conflict, so we should interpret most instances of diamonds as having that effect. While this is not as helpful for explaining the underlying causes of each conflict, it can be more useful in finding the anomalies that are conflict-prone.

As a final note, Buhaug and Rød do not appear to provide any analysis of what kind of relationship areas facing government conflict have with areas facing territorial conflict – if one can lead to the other, for instance – but it could help explain some of the findings in this study. If there is a connection between the existence of government wars and territorial wars – if instability at the center increases the chance of secessionist movements to begin – then it is also possible that there are interaction variables that this replication leaves out that ought to be included in the main study.

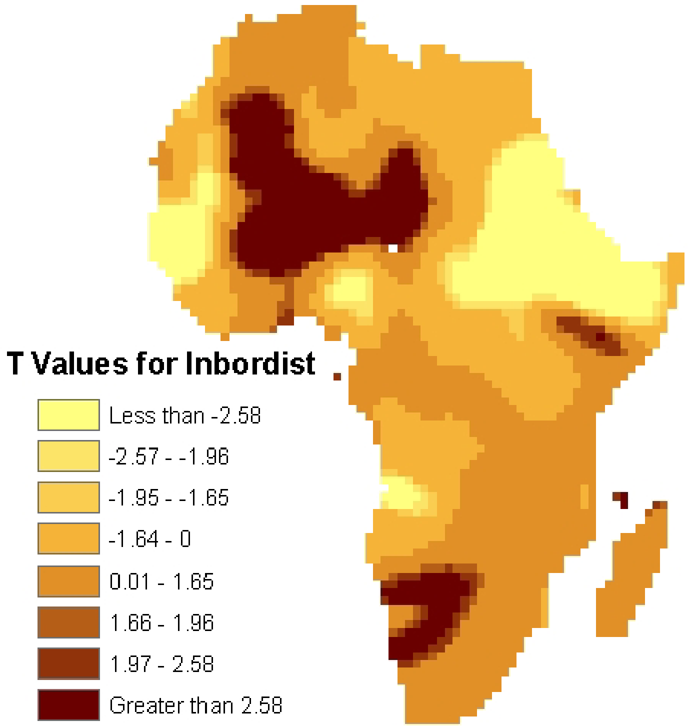

Predicted Likelihood of Territorial Conflict for Resource Model
based on GWR of Buhaug-Rod dataset
(3.4 bandwidth)
Predicted Likelihood (as Linear function)
-0.30 - -0.1
-0.09 - -0.05
-0.04 - 0.1
0.11 - 0.25
0.26 - 0.5
0.51 - 0.75
0.76 - 1
1.01 - 1.25
1.26 - 1.59


Residuals of Territorial Conflict for Resource Model
based on GWR of Buhaug-Rod dataset
(3.4 bandwidth)
Residuals
-0.77 - -0.12
-0.11 - -0.04
-0.03 - -0.01
0.00 - 0.00
0.01 - 0.00
0.01 - 0.02
0.03 - 0.07
0.08 - 0.22
0.23 - 0.89

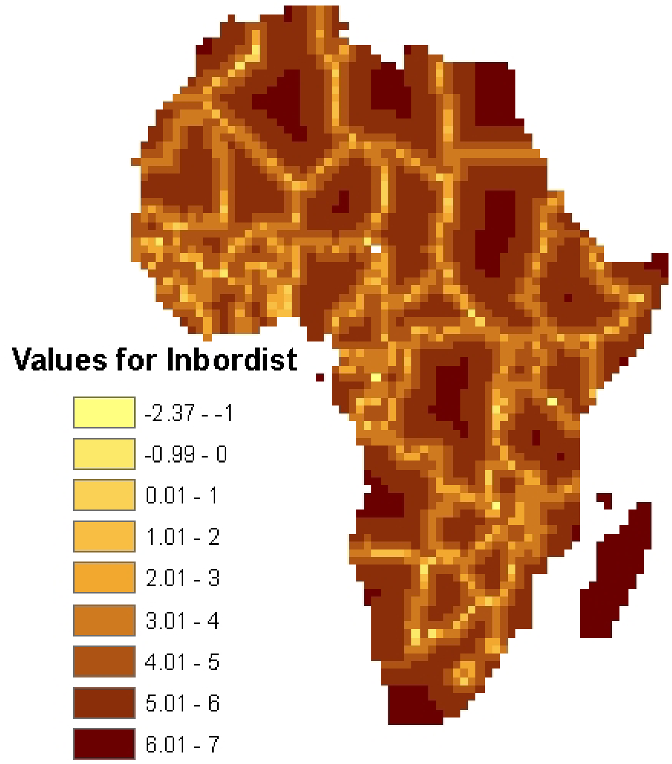

Predicted Values for Territorial Conflict
from Resource Model for Territorial Conflict
based on Logit Regression of Buhaug-Rod dataset
Predicted Values
-10.94 - -8
-7.99 - -6
-5.99 - -5
-4.99 - -4
-3.99 - -3
-2.99 - -2
-1.99 - -1
-0.99 - 0
0.01 - 2


Residuals for Territorial Conflict
from Resource Model for Territorial Conflict
based on Logit Regression of Buhaug-Rod dataset
Residual Values
-10.94 - -8
-7.99 - -6
-5.99 - -5
-4.99 - -4
-3.99 - -3
-2.99 - -2
-1.99 - -1
-0.99 - 0
0.01 - 2

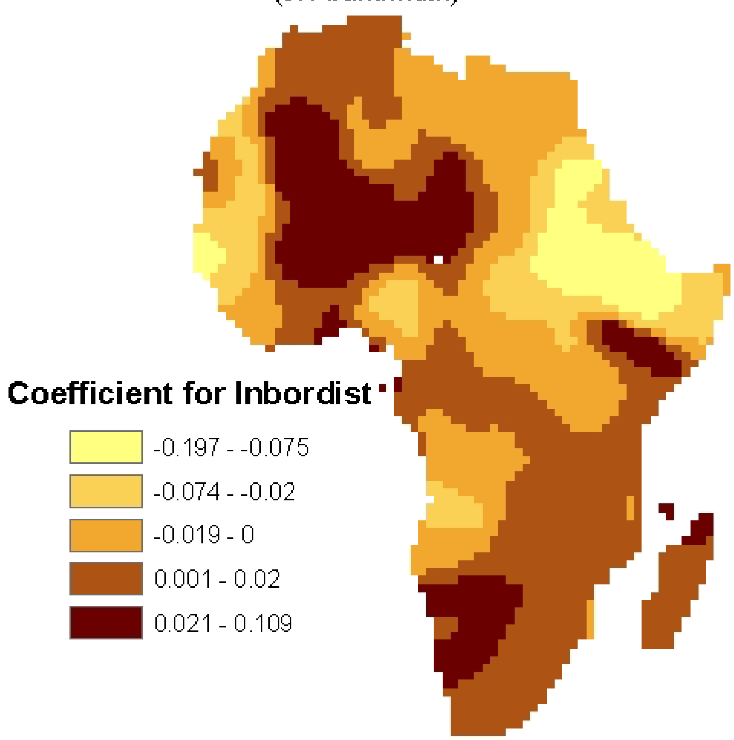
T Values for Ln Distance to Border
Resource Model for Territorial Conflict
based on GWR of Buhaug-Rod dataset
(3.4 bandwidth)

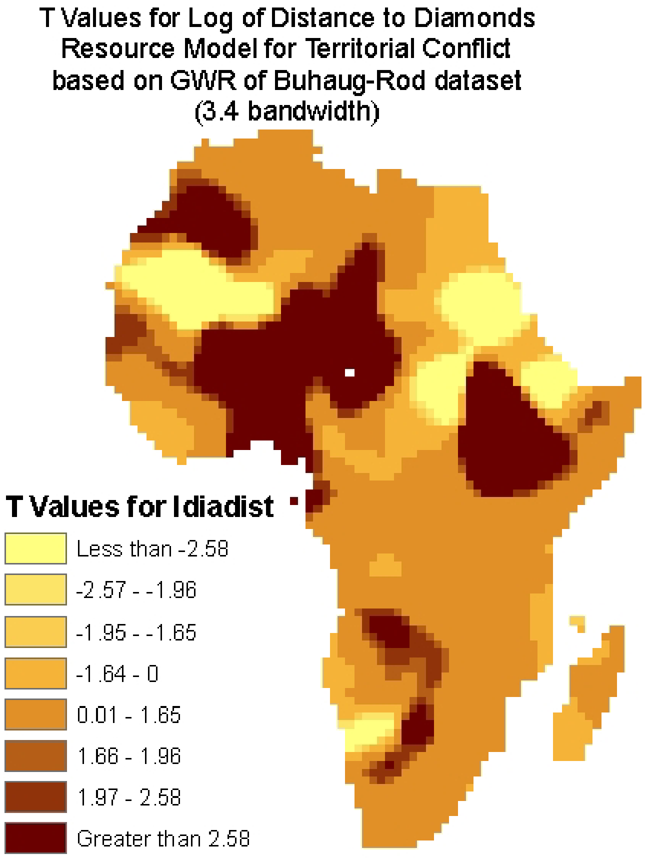


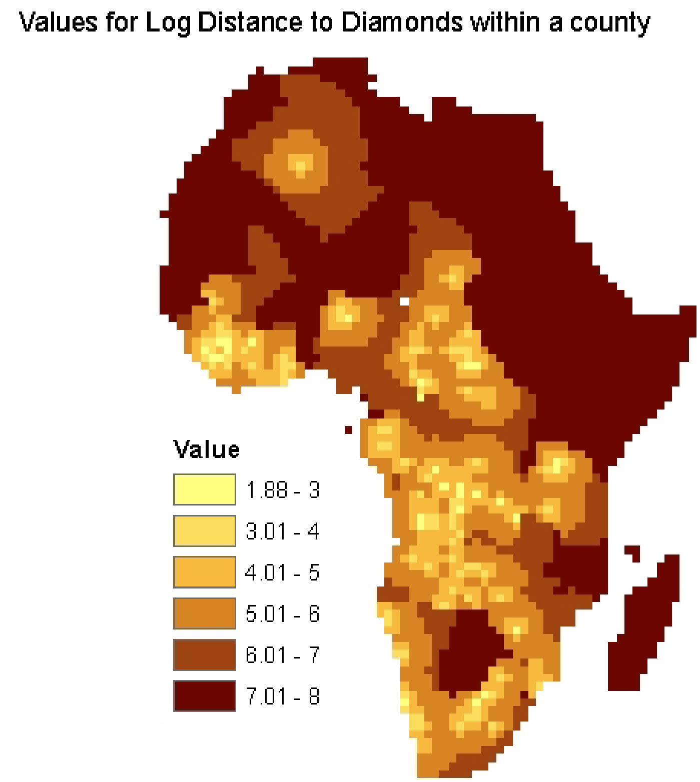
Values for Ln Distance to Border

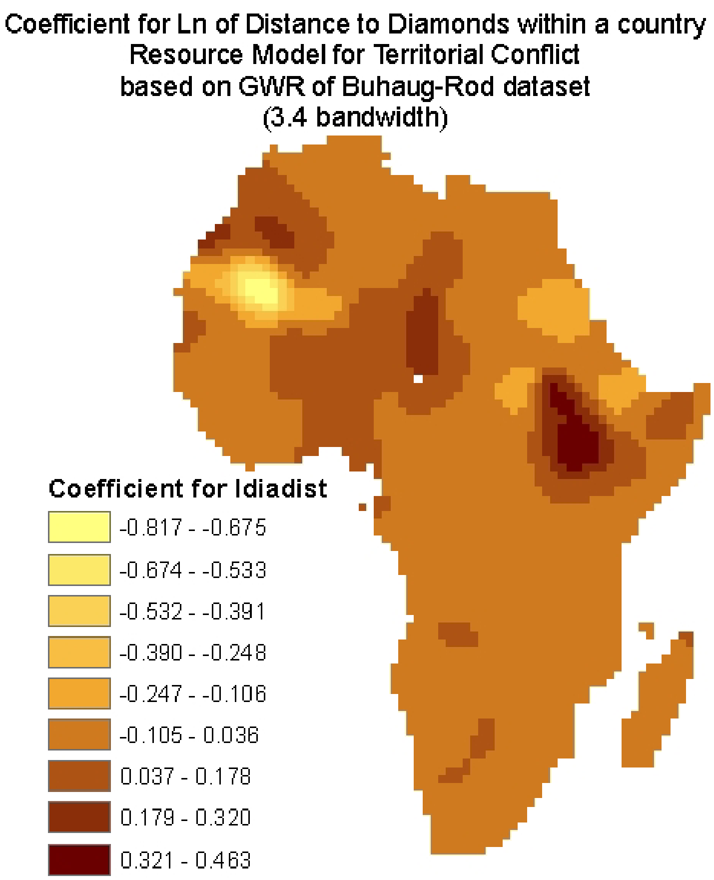


Coefficient for Ln Distance to Border
Resource Model for Territorial Conflict
based on GWR of Buhaug-Rod dataset
(3.4 bandwidth)

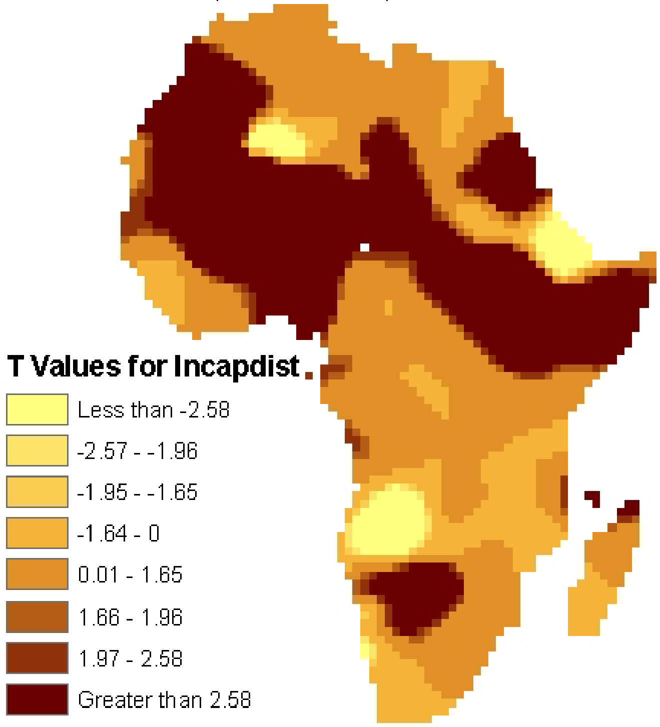



T Values for Log of Distance to Diamonds
Resource Model for Territorial Conflict
based on GWR of Buhaug-Rod dataset
(3.4 bandwidth)
T Values for ldiadist
Less than -2.58
-2.57 - -1.96
-1.95 - -1.65
-1.64 - 0
0.01 - 1.65
1.66 - 1.96
1.97 - 2.58
Greater than 2.58

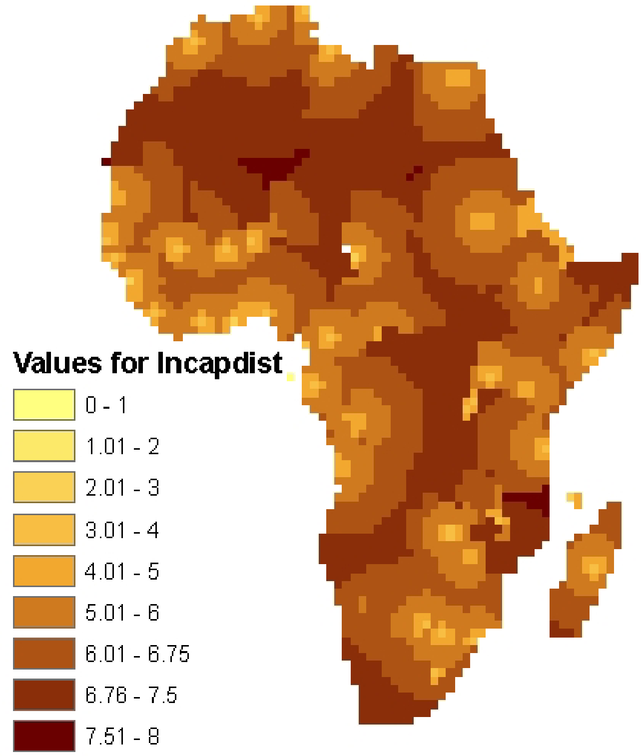

Values for Log Distance to Diamonds within a county
Value
1.88 - 3
3.01 - 4
4.01 - 5
5.01 - 6
6.01 - 7
7.01 - 8


Coefficient for Ln of Distance to Diamonds within a country
Resource Model for Territorial Conflict
based on GWR of Buhaug-Rod dataset
(3.4 bandwidth)
Coefficient for ldiadist
-0.817 - -0.675
-0.674 - -0.533
-0.532 - -0.391
-0.390 - -0.248
-0.247 - -0.106
-0.105 - 0.036
0.037 - 0.178
0.179 - 0.320
0.321 - 0.463

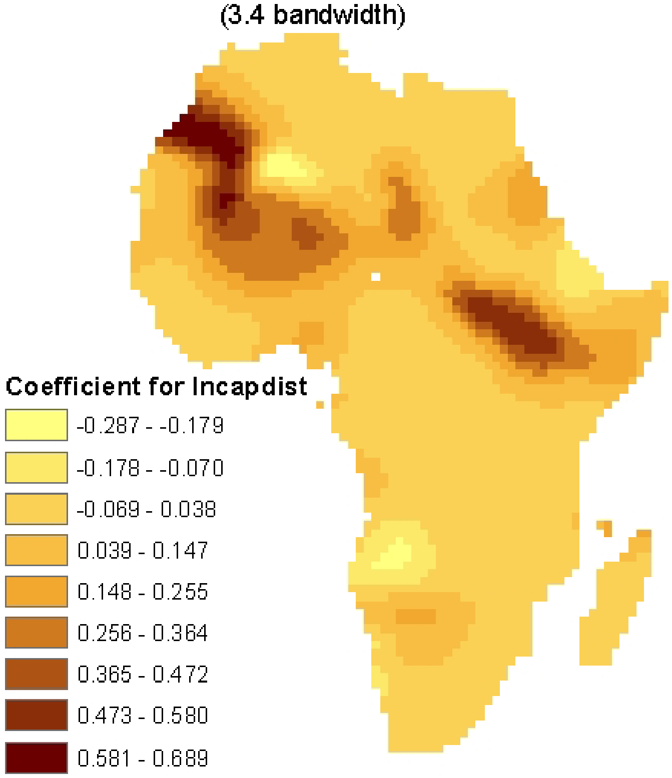
T Values for Ln of Distance to Capital within a country
Resource Model for Territorial Conflict
based on GWR of Buhaug-Rod dataset
(3.4 bandwidth)



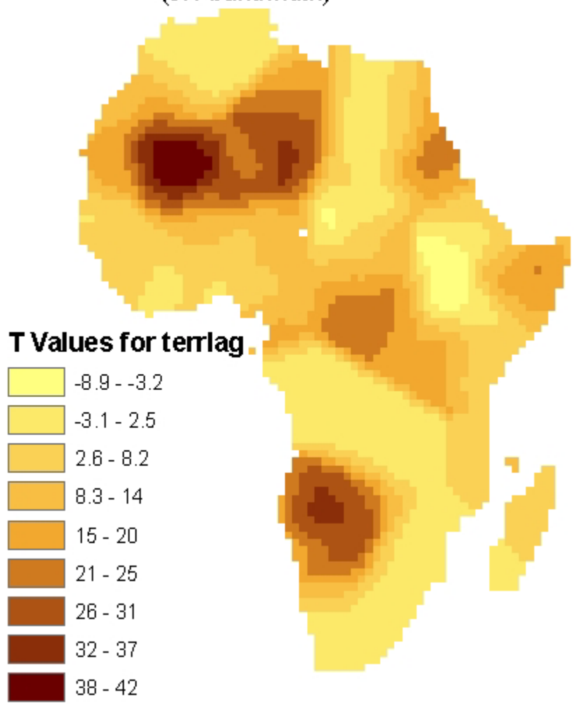
Values for Ln Distance to Capital



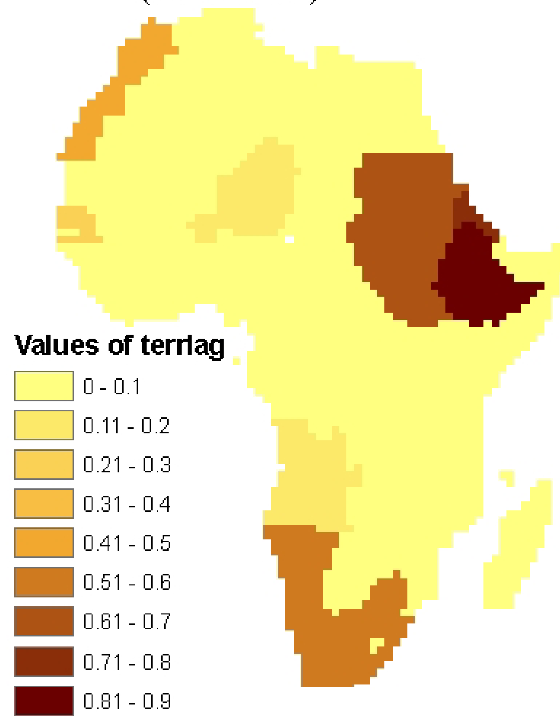
Coefficient for Ln of Distance to Capital within a country
Resource Model for Territorial Conflict
based on GWR of Buhaug-Rod dataset
(3.4 bandwidth)



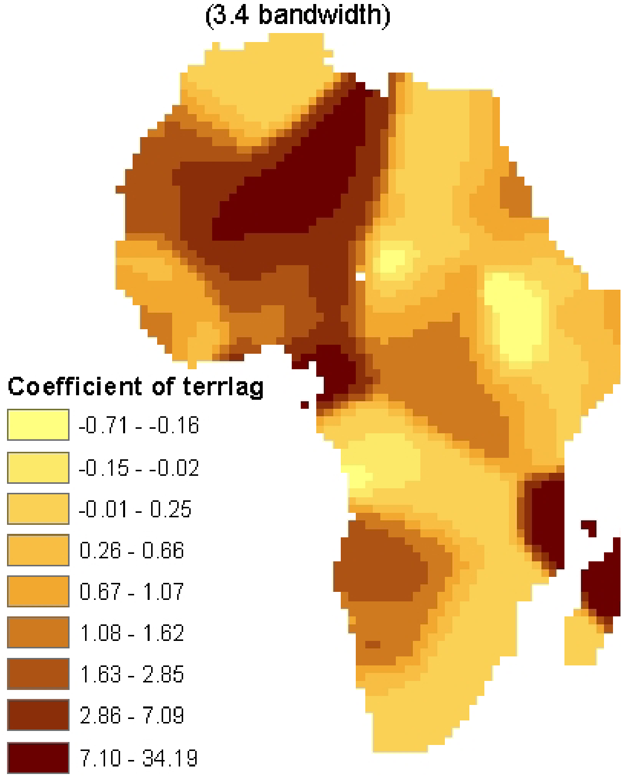
T Values for Spatial Lag to Conflict
Resource Model for Territorial Conflict
based on GWR of Buhaug-Rod dataset
(3.4 bandwidth)



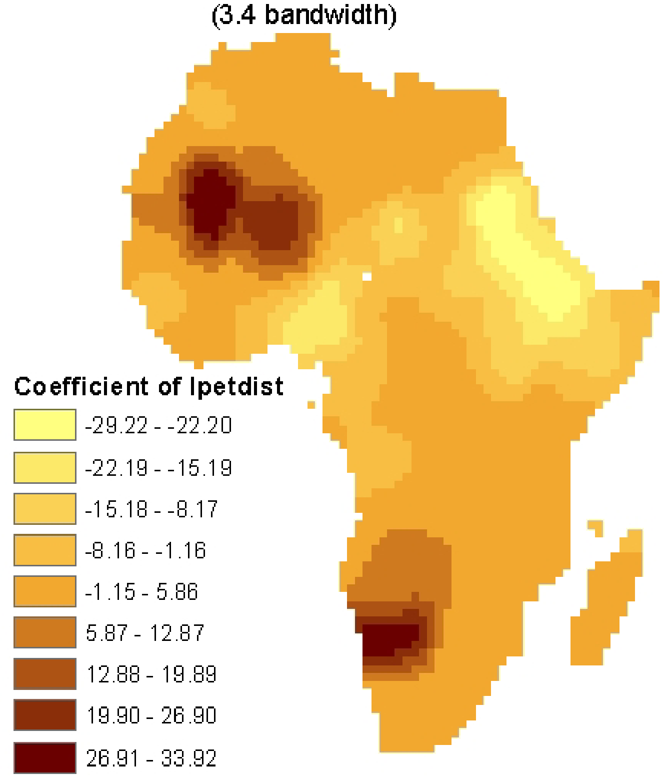
Values for Spatial Lag to Conflict
Resource Model for Territorial Conflict
based on GWR of Buhaug-Rod dataset
(3.4 bandwidth)



Coefficient for Spatial Lag to Conflict
Resource Model for Territorial Conflict
based on GWR of Buhaug-Rod dataset
(3.4 bandwidth)



Coefficient for Ln of Distance to Petroleum within a country
Resource Model for Territorial Conflict
based on GWR of Buhaug-Rod dataset
(3.4 bandwidth)



This concludes the maps for Annex 1. This page left intentionally blank.

ANNEX 2. PROPOSED VARIABLES TESTED FOR ONSET AND OCCURRENCE OF CIVIL CONFLICT

Note: Due to some changes in the data in the final stages of my thesis, the results in these tables may not completely correspond with the results from the final data file. This annex is provided as an illustrative example of the extensive process that I used to determine which variables and specifications to include in the final regression analysis.

In this section, through a modified logit regression, a series of models for conflict will be tested for their significance and effect on civil conflict onset and occurrence. Based on the experiences with the replication of the Buhaug-Rød data using geographically weighted regression (GWR), its use will be tempered in the models proposed. The most attractive element of geographically weighted regression was its ability to assess a range of variables against their geographic location. A similar effect can be achieved – though slightly more laboriously – by using a set of interaction terms to examine the specific geographic effects that are hypothesized. For instance, rather than seeing whether the impact of diamonds becomes less likely to lead to conflict near a capital and more likely near a border by examining the maps produced through GWR, these models will employ interaction terms for the presence of diamonds and the distance to state capitals or international borders.

In considering different mathematical approaches to the data, it was important to ensure that the statistical method used accurately reflected the different levels at which data was collected. Because some data – particularly the economic and political data – is collected at the country level, it was not sampled the nearly 16,000 times that it appears in the dataset for each of the six five-year periods. Instead, it was sampled about 50 times – one for each country – for each of the six five-year periods. Other data – particularly geographic data such as roughness of terrain, amount of rainfall, and even population density – were derived from rasters with even finer resolution than the 0.4 x 0.4 degree grid, and so they were sampled nearly 16,000 separate times (once for each grid cell) for each of the six five-year periods. In combining these two levels of data – the country level and the grid level – there are, broadly speaking, three approaches that could be taken – complete pooling, no pooling, and partial pooling – as laid out by Gelman and Hill in their guide to multilevel analysis (Gelman and Hill 2006). Complete pooling would entail combining all of the data in one "pool," as though the country level data had been sampled uniquely for each separate grid cell. This approach would generate less variation than is appropriate, making political and economic results appear more significant than they are. No pooling would entail a separate analysis for each country. This approach would not only eliminate all of the significant political and economic variables and prove difficult in smaller countries with very few grid cells, but it would also likely result in wildly different results for the significance and impact of geographic variables in each country, making it hard to reach broader conclusions about geographic effects. The middle ground is partial pooling, in which data from the various countries are included in the regression but are given a partial weight to reflect that the sampling for the country is unlikely to reflect perfectly the true value of the grid cell but does have some relationship. So, the data is included but weighted to have partial influence on the results.

Descriptive Statistics

With a large array of variables and significant collection of data to assess, a series of thematic examinations are arranged to test a set of variables for their relationship with conflict. The reasons for taking this approach to winnowing and constructing the models is two-fold. First, many of these variables are viewed as theoretically similar (even if they are not always highly correlated with each other), and in a general model, it would be helpful to have only one or two variables representing a strand of theoretical argument. Second, many policymakers and analysts are most familiar with one type of information – geographers know terrain, economists know income metrics, demographers know population trends, and so forth. While these professions certainly are not stove-piped into understanding only the information most familiar to them, it is regularly a first source for information. Thus, it would be useful to know which information, within a set of variables commonly tracked and observed by different groups, are the most pertinent to civil conflict.

The first move with data analysis is to establish a handful of thematic models. In this analysis, many of these variables are partially correlated – this situation is understandable since the variables often come from the same set of sources. In some cases, such as political data, the sources are even component scores of the larger composite scores in Polity IV. The advantage of this approach is that it allows more significant variables in a set to be identified as a general model is built. Also, with dozens of variables and as many as six dependent variables (for government and territorial conflict onset, occurrence, and battle deaths), it will make the process more digestible to examine sets of variables at a time rather than over fifty at once.

Before building full models, as will happen in the next chapter, though, it is helpful to examine a simple table of overall statistics comparing the means of the conflict and non-conflict grid cells. This series of tables gives a general sense of which variables are likely to have the most significant role and which are likely to have a much less significant influence, and it explains the background for the selection of a subset of variables for the final model. It is important to note that any one of these variables with a significant difference between their conflict and non-conflict cells could ultimately prove to be inconsequential in the regression models when other variables are included as controls.

It is also worth reiterating why multiple specifications of the same kinds of variables (economic, rainfall, terrain, political scores, etc) are included. In general, while there are conceptual frameworks that lead researchers to use certain variables – rainfall as an exogenous measure of economic growth, for instance – they do not always lead to clear ideas about what specifications of a variable are appropriate. Just because there is a hypothesis that lower amounts of rainfall should lead to conflict does not specify how those lower amounts should be assessed – in absolute terms, in relative decrease, per capita, etc. To make these distinctions, the project relies on the data. As distinctions are made between different specifications of a variable, it should lead to a clearer understanding of what aspects of the hypothesis hold the greatest weight. For instance, if absolute rainfall is important, but changes in rainfall appear less important, then one might conclude that conflict is simply most common in areas without many resources for agriculture or growing a basic economy. On the other hand, if absolute rainfall is irrelevant, but changes in rainfall are important, then one might conclude that it is primarily the shocks to the system that make a region likely to experience conflict, and that social structures and practices have been established to peacefully handle the lack of rain.

In this section the short codes of the variable names are used, both for reasons of space and ease. Below is a table that connects the short codes of the variables to their full variable names.

Table of variable names

meanrain	average amount of rain in a grid cell over a 5-year period
raingrow	% increase in the amount of rain in a grid cell compared to the previous 5-year period
raindiff	absolute difference in the amount of rain in a grid cell compared to the previous 5-year period
meanrainpc	average amount of per capita rain in a grid cell
raindiffpc	difference in the per capital amount of rain in a grid cell compared to the previous 5-year period
cow_irst	country iron and steel production data from the COW dataset
cow_energy	country energy consumption data from the COW dataset

gg_lngdp_	natural log of the 5-year average GDP of a country, from the CH dataset
gg_gy1	growth in GDP of a country compared to the previous 5-year period, from the CH dataset
gg_ygini	country income Gini factor, from the CH dataset
gg_lgini	country land ownership Gini factor, from the CH dataset
unuwider_gini	country income Gini factor data from UNU-WIDER dataset
meanrains~le	difference between average rainfall in a grid cell and the average rainfall in a country
elevation	the average elevation in the grid cell (also called "Local Elevation")
elevrange	the range of elevation in the grid cell (also called "Local Elevation Range")
elevslopep~2	the percentage slope in the grid cell (also called "Local Elevation Slope as %")
elevationh~s	a scale where values are 1 minus the percentile elevation difference from the 50th percentile (also called "Local Elevation Hills")
elevrangeh~s	a scale where values are 1 minus the percentile range difference from the 50th percentile (also called "Hills Elevation Range")
elevs~2hills	a scale where values are 1 minus the percentile slope difference from the 50th percentile (also called "Hills Slope as %")
forestcover	a scale from 0 to 3 for the density of forests (also called "Local Forest Density Scale")

forestmid	a scale where values are 1 minus the percentile density difference from the 50th percentile (also called "Local Forest Middle")
forestextr~e	a scale where values are 1 minus the percentile elevation difference from the 0th or 100th percentile (also called "Local Forest Extreme")
capdist	straight-line distance from the grid cell to the state capital
cappathdist	distance from the grid cell to the state capital taking into account elevation
borddist	straight-line distance from the grid cell to an international border
bordpathdist	distance from the grid cell to an international border taking into account elevation
capborddist	ratio of the straight-line distances from the grid cell to the state capital and to an international border
capbordpat~t	ratio of the distances from the grid cell to the state capital and to an international border taking into account elevation
poplocal	local population in a grid cell
poplnlocal	natural log of the local population in a grid cell
cow_tpop	total state population from COW dataset
cow_upop	total urban population in a state from COW dataset
popdensest~e	population density in a grid cell
popratelocal	population growth rate for a grid cell compared with previous 5 year period

popratestate	population growth rate for a state compared with previous 5 year period
gg_sxp	the percentage of primary commodity exports in a country's economy from the CH dataset
gg_sxp2	the squared percentage of primary commodity exports in a country's economy from the CH dataset
gg_oilsxp	the percentage of oil commodity exports in a country's economy from the CH dataset
gg_oilsxp2	the squared percentage of oil commodity exports in a country's economy from the CH dataset
petrodist	the distance from a grid cell to a petroleum deposit
petrodense	the density of petroleum deposits around a grid cell
mineralsde~e	the density of mineral deposits around a grid cell
diasecdens~y	the density of secondary diamond deposits around a grid cell
manganesed~e	the density of manganese deposits around a grid cell
golddense	the density of gold deposits around a grid cell
copperdense	the density of copper deposits around a grid cell
platinumde~e	the density of platinum deposits around a grid cell
chromiumde~e	the density of chromium deposits around a grid cell
coltandense	the density of coltan deposits around a grid cell
irondense	the density of iron deposits around a grid cell

`dialdense`	`the density of lootable diamond deposits around a grid cell`
`dialypdense`	`the density of lootable diamond deposits with known production around a grid cell`
`dianldense`	`the density of non-lootable diamond deposits around a grid cell`
`dianlypdense`	`the density of non-lootable diamond deposits with known production around a grid cell`
`diasecdist`	`the distance from a grid cell to a secondary diamond deposit`
`p4_democ_5y`	`the average democracy score in Polity IV for the 5-year-period`
`p4_democ_t~x`	`the adjusted democracy score in Polity IV for the 5-year-period (see section on Data Sources for method)`
`p4_autoc_5y`	`the average autocracy score in Polity IV for the 5-year-period`
`p4_autoc_t~x`	`the adjusted autocracy score in Polity IV for the 5-year-period (see section on Data Sources for method)`
`p4_polity2~s`	`the average overall score in Polity IV for the 5-year-period`

Economic and Rainfall Data

One idea reviewed earlier in this thesis was the concept that economics played a key role in determining conflict patterns, and, while economic data might not be available on a local level, rainfall data could be an acceptable proxy. Under this theory, one would expect to see some level of correlation between state income and rainfall patterns. Below shows a comprehensive table of correlations for variable under consideration in this process.

Correlation Table for Economic and Rainfall Data

(obs=84507)

	meanrain	raingrow	raindiff	meanra~c	raindi~c	cow_irst	cow_en~y	gg_lng~_	gg_gy1
meanrain	1.0000								
raingrow	-0.0526	1.0000							
raindiff	-0.0152	0.3420	1.0000						
meanrainpc	0.2471	-0.0080	0.0047	1.0000					
raindiffpc	0.0119	0.2670	0.4125	-0.0328	1.0000				
cow_irst	-0.1449	0.0361	0.0418	-0.0501	0.0201	1.0000			
cow_energy	-0.2653	0.0447	0.0335	-0.0899	0.0254	0.5874	1.0000		
gg_lngdp_	-0.1002	-0.0027	-0.0368	0.0111	-0.0191	0.3106	0.3956	1.0000	
gg_gy1	-0.1998	0.0223	-0.0289	-0.0267	-0.0344	0.0513	0.0808	0.2044	1.0000

From this table, it is clear that the local rainfall data has little relationship with the statewide economic information. Unsurprisingly, the rainfall data has its highest correlations with related variables, so the absolute rain difference between the current period and the period prior has the highest correlation on the table with rain difference per capita. Meanwhile, the highest correlation between state and local data is the correlation between the mean amount of annual rainfall and the economic growth in GDP. It is interesting to note that the two measures for economic growth that from the Correlates of

War project – one for energy consumption and the other for iron and steel production – do not have a particularly high correlation with Collier and Hoeffler's data for the GDP of the countries under observation.

Relation to Onset of Conflict

The tables below show the connection between government conflict onset and economic data. The data for this section was separated into the rainfall data, which has been collected at the grid cell level, and the aggregate economic data, which is collected at the state level. The aggregate economic data has its standard errors and confidence intervals clustered on the country-year level to provide a more reasonable assessment.

It is interesting that absolute mean annual rainfall appears to have a strong relationship with governmental conflict. Given the limited changes in the weather patterns when considered on a large scale, this data – if it holds up during regression with other control variables – would be a surprising find. It would be particularly surprising given that initial hypotheses were that rainfall would be a conflict deterrent. While that appears to be true in the case of territorial conflict, more rainfall appears to have an association with higher governmental conflict onset. Given that this study focuses on Africa, the Sahara may be playing a major role in associating lower rainfall with territorial conflict, and the Congo may also be playing a role as its large nationwide governmental conflict has generally high rainfall figures. Another interesting early finding from the data is that the rain growth from the previous five-year period appears to be positive associated with both territorial and governmental conflict, with a particularly strong distinction in the territorial conflict occurrence.

With the use of other control variables in the model, mean rainfall could become less significant, but this early table highlights how significant some environmental factors can be.

Mean Estimates for Rainfall Data and Onset of Government Conflict

```
Mean estimation                         Number of obs    =   95886

            0: govonset = 0
            1: govonset = 1

--------------------------------------------------------------
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
meanrain     |
           0 |   1053.255   3.028134      1047.319     1059.19
           1 |   1753.574   15.83466      1722.538     1784.61
-------------+------------------------------------------------
raingrow     |
           0 |   .0201644   .0010461      .0181141    .0222148
           1 |   .0315819   .0016103      .0284257    .0347382
-------------+------------------------------------------------
raindiff     |
           0 |  -10.40579   .4119975      -11.2133   -9.598279
```

```
           1 |   51.89497    2.27703      47.43201    56.35792
-------------+------------------------------------------------
meanrainpc   |
           0 |   117.8464   .9320311      116.0197    119.6732
           1 |   162.9711   5.723393      151.7534    174.1889
-------------+------------------------------------------------
raindiffpc   |
           0 |  -1.219979   .0992414     -1.414491   -1.025467
           1 |   5.131484   .5668557      4.020454    6.242515
--------------------------------------------------------------
```

Mean Estimates for Rainfall Data and Onset of Territorial Conflict

```
Mean estimation                    Number of obs    =   95886

            0: teronset = 0
            1: teronset = 1

--------------------------------------------------------------
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
meanrain     |
           0 |   1083.562   3.038614      1077.607    1089.518
           1 |   625.9516   13.67446      599.1498    652.7534
-------------+------------------------------------------------
raingrow     |
           0 |    .019434   .0010284      .0174183    .0214496
           1 |   .0814915   .0048236      .0720373    .0909456
```

raindiff					
	0	**-8.582654**	**.4111827**	**-9.388568**	**-7.776741**
	1	**.2716883**	**2.795266**	**-5.207002**	**5.750379**
meanrainpc					
	0	**119.7839**	**.9352629**	**117.9508**	**121.6171**
	1	**91.16888**	**3.476039**	**84.35589**	**97.98188**
raindiffpc					
	0	**-1.054306**	**.0990133**	**-1.248371**	**-.8602412**
	1	**.9947197**	**.5904301**	**-.1625167**	**2.151956**

Of the four economic variables assessed, only growth in GDP appears to have a possible relationship with governmental conflict. Given the strong literature suggesting that absolute GDP also has a significant relationship, this result is surprising. (When this mean estimation is performed without clustering the errors, the finding for absolute GDP is significant.) It is also interesting that the economic data appears to have less significance when assessed against territorial conflict. The only variable that appears significant in that mean estimation is the Correlate of War measure for iron and steel production. This finding is also surprising, since it would be expected that both growth in GDP and absolute GDP would be significant based on aggregate results from earlier studies. It may be that more control variables are necessary, such as distance from the capital, to account for some of the variance resulting from measuring in grid cells rather than at the state level.

Mean Estimates for Economic Data and Onset of Government Conflict

```
Mean estimation                         Number of obs    =   84507

            0: govonset = 0
            1: govonset = 1

               (Std. Err. adjusted for 267 clusters in c_year)
--------------------------------------------------------------
             |               Robust
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
cow_irst     |
           0 |   532.5632   164.1282      209.4075    855.7188
           1 |   125.9054   106.6077     -83.99693    335.8076
-------------+------------------------------------------------
cow_energy   |
           0 |   22828.63   4690.905       13592.6    32064.65
           1 |   30229.55   24645.69      -18295.9       78755
-------------+------------------------------------------------
gg_lngdp_    |
           0 |   6.700049   .1310022      6.442116    6.957982
           1 |   6.225519   .4061166      5.425907    7.025131
-------------+------------------------------------------------
gg_gy1       |
           0 |   .1750936   .2534083     -.3238477    .6740349
           1 |  -6.573055   1.887414     -10.28923   -2.856884
--------------------------------------------------------------
```

Mean Estimates for Economic Data and Onset of Territorial Conflict

```
Mean estimation                     Number of obs    =   84507

            0: teronset = 0
            1: teronset = 1

               (Std. Err. adjusted for 267 clusters in c_year)
--------------------------------------------------------------
             |               Robust
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
cow_irst     |
           0 |   528.4667   161.5266      210.4334       846.5
           1 |   .2870337   .2772416     -.2588334    .8329009
-------------+------------------------------------------------
cow_energy   |
           0 |   23497.48   5011.735      13629.76    33365.19
           1 |    3375.13    1204.53      1003.505    5746.755
-------------+------------------------------------------------
gg_lngdp_    |
           0 |   6.688742   .1308625      6.431084      6.9464
           1 |   6.398958   .1572358      6.089373    6.708543
-------------+------------------------------------------------
gg_gy1       |
           0 |  -.0716501   .3000479     -.6624212    .5191209
           1 |   .1080166   .8395543     -1.545001    1.761034
--------------------------------------------------------------
```

Relation to Occurrence of Conflict

Mean Estimates for Rainfall Data and Occurrence of Governmental Conflict

```
Mean estimation                     Number of obs    =   95886

            0: govonset = 0
            1: govonset = 1

--------------------------------------------------------------
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
meanrain     |
           0 |   1053.255   3.028134      1047.319     1059.19
           1 |   1753.574   15.83466      1722.538     1784.61
-------------+------------------------------------------------
raingrow     |
           0 |   .0201644   .0010461      .0181141    .0222148
           1 |   .0315819   .0016103      .0284257    .0347382
-------------+------------------------------------------------
raindiff     |
           0 |  -10.40579   .4119975      -11.2133   -9.598279
           1 |   51.89497    2.27703      47.43201    56.35792
-------------+------------------------------------------------
meanrainpc   |
           0 |   117.8464   .9320311      116.0197    119.6732
           1 |   162.9711   5.723393      151.7534    174.1889
```

raindiffpc				
0	**-1.219979**	**.0992414**	**-1.414491**	**-1.025467**
1	**5.131484**	**.5668557**	**4.020454**	**6.242515**

Mean Estimates for Rainfall Data and Occurrence of Territorial Conflict

Mean estimation Number of obs = 95886

0: teronset = 0

1: teronset = 1

Over	Mean	Std. Err.	[95% Conf. Interval]	
meanrain				
0	**1083.562**	**3.038614**	**1077.607**	**1089.518**
1	**625.9516**	**13.67446**	**599.1498**	**652.7534**
raingrow				
0	**.019434**	**.0010284**	**.0174183**	**.0214496**
1	**.0814915**	**.0048236**	**.0720373**	**.0909456**
raindiff				
0	-8.582654	.4111827	-9.388568	-7.776741
1	.2716883	2.795266	-5.207002	5.750379

meanrainpc					
	0	**119.7839**	**.9352629**	**117.9508**	**121.6171**
	1	**91.16888**	**3.476039**	**84.35589**	**97.98188**
raindiffpc					
	0	-1.054306	.0990133	-1.248371	-.8602412
	1	.9947197	.5904301	-.1625167	2.151956

Horizontal Inequality

For horizontal inequality, there are fewer variables available to provide measures, and they are almost all at the state level. This section has the shakiest data both in terms of reliability and quantity, but given the strong theoretical arguments presented for its importance, an effort to include horizontal inequality cannot be dismissed. The variables available at the aggregate level are gathered from the Collier and Hoeffler study including measures for the Gini factor in income and land ownership and from the UN University WIDER project. At the disaggregate level, variables are more sparse, but a measure that compares the amount of mean rain in a grid cell against the country average has been used. In total, there are only about 32,000 grid cells, about one-third of the total, where data on all three inequality measures is present. And, as can be seen in the correlation table below, it is concerning that the three Gini factors – two for income and one for land – are not more heavily correlated. It suggests that the methods of collecting the data and the groups that were included in the survey are so distinct that this data may prove not useful. The reason that the rainfall data has almost exactly zero correlation with the aggregate variables is that it measures the difference in rainfall in a local grid cell compared to the national average for a country in a given five-year period. As a result, the values are in general spread evenly around the mean for the state, while the other three variables are constant across the entire country-period.

Correlation Table for Economic and Rainfall Inequality Data

```
(obs=32244)

             | gg_ygini gg_lgini unuwid~i meanr~le
-------------+------------------------------------
    gg_ygini |   1.0000
    gg_lgini |   0.2420   1.0000
unuwider_g~i |   0.0046  -0.1230   1.0000
meanrains~le |  -0.0000  -0.0000   0.0000   1.0000
```

Relation to Onset of Conflict

Because of the vast differences in the amount of coverage that each of these variables has across the dataset, each data source has a separate mean estimation table, rather than performing them in one single analysis. Remarkably, the data below indicates that despite having little correlation with each other, the Gini factors for the countries evaluated, from the both the Collier-Hoeffler data and the UNU WIDER data, return different means with greater than 95% confidence for government conflict onset. Meanwhile, land inequality data and the difference in rainfall from the state average proved less compelling, and none of the data on inequality provided a strong explanation of territorial conflict. It is tempting to plumb the reasons why the value of the variable representing the rainfall difference from the mean did not have greater significance, perhaps because extremely low and high levels of rainfall are more conducive to conflict. But, since the extremes of rainfall inequality are closely associated with the extremes of forest cover, for which there is a much stronger theoretical justification for examining the extreme values, the study does push to find a strong specification for rainfall inequality.

Mean Estimates for Collier-Hoeffler Inequality Data and Onset of Governmental Conflict

```
Mean estimation                     Number of obs    =   87402

            0: govonset = 0
            1: govonset = 1

               (Std. Err. adjusted for 272 clusters in c_year)
--------------------------------------------------------------
             |               Robust
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
gg_ygini     |
           0 |   25.54745   1.863845      21.87799    29.21691
           1 |   7.911928   5.600691     -3.114469    18.93832
-------------+------------------------------------------------
gg_lgini     |
           0 |   .2810052   .0262723      .2292813     .332729
           1 |   .2858141   .1432494      .0037909    .5678372
--------------------------------------------------------------
```

Mean Estimates for UNU-WIDER Inequality Data and Onset of Governmental Conflict

```
Mean estimation                    Number of obs    =   33710

            0: govonset = 0
            1: govonset = 1

               (Std. Err. adjusted for 107 clusters in c_year)
--------------------------------------------------------------
             |               Robust
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
unuwider_g~i |
           0 |   48.17538   1.194339      45.80748    50.54327
           1 |   38.14912   2.984296      32.23247    44.06578
--------------------------------------------------------------
```

Mean Estimates for Rainfall Inequality Data and Onset of Governmental Conflict

```
Mean estimation                    Number of obs    =   95886

            0: govonset = 0
            1: govonset = 1

               (Std. Err. adjusted for 302 clusters in c_year)
--------------------------------------------------------------
             |               Robust
```

```
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
meanrains~le |
           0 |  -1.401886   1.899411     -5.139693    2.335921
           1 |   42.72853   56.53978     -68.53478    153.9918
--------------------------------------------------------------
```

Mean Estimates for Collier-Hoeffler Inequality Data and Onset of Territorial Conflict

```
Mean estimation                     Number of obs    =   87402

            0: teronset = 0
            1: teronset = 1

               (Std. Err. adjusted for 272 clusters in c_year)
--------------------------------------------------------------
             |               Robust
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
gg_ygini     |
           0 |    24.9747   1.863278      21.30636    28.64304
           1 |   22.81905   8.887616      5.321503     40.3166
-------------+------------------------------------------------
gg_lgini     |
           0 |   .2781421   .0262594      .2264438    .3298404
           1 |   .4349679   .0722726      .2926807    .5772551
--------------------------------------------------------------
```

Mean Estimates for UNU-WIDER Inequality Data and Onset of Territorial Conflict

```
Mean estimation                     Number of obs    =   33710

            0: teronset = 0
            1: teronset = 1

               (Std. Err. adjusted for 107 clusters in c_year)
--------------------------------------------------------------
             |               Robust
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
unuwider_g~i |
           0 |   48.01048    1.20013       45.6311    50.38985
           1 |   52.30988   4.430657      43.52568    61.09409
--------------------------------------------------------------
```

Mean Estimates for Rainfall Inequality Data and Onset of Territorial Conflict

```
Mean estimation                     Number of obs    =   95886

            0: teronset = 0
            1: teronset = 1

               (Std. Err. adjusted for 302 clusters in c_year)
--------------------------------------------------------------
             |               Robust
```

```
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
meanrains~le |
           0 |    .397563   3.916943     -7.310497    8.105623
           1 |  -22.17243   219.7846     -454.6815    410.3366
--------------------------------------------------------------
```

Relation to Occurrence of Conflict

Similar to the data on onset, the aggregate and disaggregate data on occurrence is not particularly compelling. As with onset, the only areas of possible significance in aggregate measures for inequality relating to governance conflicts, but in this case, the Collier-Hoeffler data on both income and land inequality is all that appears significant.

Mean Estimates for Collier-Hoeffler Inequality Data and Occurrence of Governmental Conflict

```
Mean estimation                     Number of obs    =   87402

            0: govoccur = 0
            1: govoccur = 1

               (Std. Err. adjusted for 272 clusters in c_year)
--------------------------------------------------------------
             |               Robust
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
gg_ygini     |
```

```
           0 |   27.35675   1.898598      23.61887    31.09462
           1 |   10.93404   3.449028      4.143744    17.72433
-------------+--------------------------------------------------
gg_lgini     |
           0 |   .3057486    .027476       .251655    .3598421
           1 |   .1392257   .0514716      .0378907    .2405607
----------------------------------------------------------------
```

Mean Estimates for UNU-WIDER Inequality Data and Occurrence of Governmental Conflict

```
Mean estimation                     Number of obs    =   33710

            0: govoccur = 0
            1: govoccur = 1

               (Std. Err. adjusted for 107 clusters in c_year)
----------------------------------------------------------------
             |               Robust
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+--------------------------------------------------
unuwider_g~i |
           0 |   48.48095    1.19103      46.11962    50.84228
           1 |   42.97873   2.603585      37.81687    48.14059
----------------------------------------------------------------
```

Mean Estimates for Rainfall Inequality Data and Occurrence of Governmental Conflict

```
Mean estimation                    Number of obs    =   95886

            0: govoccur = 0
            1: govoccur = 1

               (Std. Err. adjusted for 302 clusters in c_year)
--------------------------------------------------------------
             |                Robust
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
meanrains~le |
           0 |   .5298524   3.400177     -6.161277    7.220982
           1 |  -3.408552   21.91613     -46.53679    39.71969
--------------------------------------------------------------
```

Mean Estimates for Collier-Hoeffler Inequality Data and Occurrence of Territorial Conflict

```
Mean estimation                    Number of obs    =   87402

            0: teroccur = 0
            1: teroccur = 1

               (Std. Err. adjusted for 272 clusters in c_year)
--------------------------------------------------------------
             |                Robust
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
gg_ygini     |
```

```
           0 |   25.30439   1.869645      21.62352    28.98527
           1 |   20.50949   4.213469       12.2142    28.80478
-------------+--------------------------------------------------
gg_lgini     |
           0 |   .2708139   .0256461      .2203229    .3213049
           1 |   .4045679   .0530305      .3001637    .5089721
--------------------------------------------------------------
```

Mean Estimates for UNU-WIDER Inequality Data and Occurrence of Territorial Conflict

```
Mean estimation                     Number of obs    =   33710

            0: teroccur = 0
            1: teroccur = 1

               (Std. Err. adjusted for 107 clusters in c_year)
--------------------------------------------------------------
             |               Robust
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
unuwider_g~i |
           0 |   48.05551   1.252005      45.57329    50.53773
           1 |   49.08283   2.096311      44.92669    53.23897
--------------------------------------------------------------
```

Mean Estimates for Rainfall Inequality Data and Occurrence of Territorial Conflict

```
Mean estimation                     Number of obs    =   95886

            0: teroccur = 0
            1: teroccur = 1

               (Std. Err. adjusted for 302 clusters in c_year)
--------------------------------------------------------------
             |              Robust
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
meanrains~le |
           0 |  -11.08061   8.663697      -28.1297    5.968474
           1 |   145.8814   103.9333     -58.64647    350.4092
--------------------------------------------------------------
```

Rough Terrain and Forests Data

In assessing elevation and rough terrain, the broad concept that this project has put forth is that it neither condition should have a direct linear relationship with conflict. Neither the highest elevation or just the densest forests – that make sense for conflict. Neither rebels nor government forces benefit from the difficult conditions of high elevations more than it costs them. Instead, it is mid-range elevations that create just enough difficulty for it to affect the ability of government to penetrate without the terrain making it to difficult to sustain rebel operations. Similarly, while dense forests do not present the resource difficulties (access to food and water) that high elevations can, it is also true that extremely low forest levels – desert – can be equally foreboding to outsiders not familiar with the terrain.

There are no particularly surprising correlations from the table below. As one would expect, the grid cells with higher elevations also have greater ranges of elevations and more severe slopes, but not to such a high degree of correlation that assessing them separately is unnecessary.

One of the other questions that this data can examine is whether the same factors that cause conflict to in a country in general determine the location within the country. For this purpose, data on the average elevation, elevation range, and slope in each country has been included. Also unsurprising, the local data and the state data on the same variable (absolute elevation, elevation range, etc) are generally somewhat correlated.

Note: because so many different specifications are used in this section, the short names have been converted to more specific terms.

Correlation Table for Elevation Data

(obs=95718)

	elevat~l	elevat~e	elevra~l	elevra~e	elevsl~l	elevsl~e	elevat~s	elevra~s	elevsl~s
Local Elevation	1.0000								
State Elevation	0.6264	1.0000							
Local ElevRange	0.0533	0.0636	1.0000						
State ElevRange	0.0169	-0.0033	0.4697	1.0000					
Local Slope %	0.0299	0.0305	0.5468	0.3482	1.0000				

State Slope % Avg	-0.0453	-0.0890	0.4517	0.9889	0.3415	1.0000			
Hills Elevation	-0.3952	-0.0165	0.2155	-0.0681	0.1018	-0.0843	1.0000		
Hills ElevRange	-0.0044	-0.0188	-0.5294	-0.0832	-0.2849	-0.0872	-0.1901	1.0000	
Hills Slope %	-0.0260	-0.0588	-0.4079	-0.1831	-0.4988	-0.1701	-0.0762	0.3421	1.0000

Correlation Table for Forest Data

(obs=95886)

	f~rlocal	forest~e	f~dlocal	f~elocal
Local Forest	1.0000			
State Forest	0.8313	1.0000		
Local Middle	-0.0282	0.0975	1.0000	
Local Extreme	-0.0290	-0.1492	-0.9835	1.0000

Relation to Onset of Conflict

The major analytical point that emerges from the analysis of the elevation and forest density variables is that while higher elevations and denser forests are generally good for conflict, it is the mid-points of the variables that carry much of the weight. For instance, contrary to some of the initial hypotheses but consistent with the overall thrust of the data, the mid-ranges for forests appear to be more prone to governmental conflict, while the extremes are less likely to face governmental conflict. The opposite appears to be true for territorial conflict, where the mid-ranges for forest density are less likely to see conflict and the extremes are more likely. More broadly, this data indicates that governmental conflict is more likely in places where groups in society are more likely to interact with the state – where the terrain is more reasonable, where the foliage is less forbidding – but territorial conflict is more likely in cases where societal groups are less likely to have contact with the state – where the elevations are less reasonable, and the conditions is less penetrable.

In reading these tables, two notes: First, it is also worth remembering that while the distinction between the variables related to hills may seem small, they are on a constructed percentile scale that ranges from 0.5 to 1, so the distinction of being on either side of 0.75 is a meaningful one. Second, the tendency for territorial conflict to occur in extreme values of forest cover may be somewhat confounded by the data from the Sahara. To have greater confidence of these results, replication in another continent, such as Asia or Latin America, would be helpful.

Mean Estimates for Local Terrain Data and Onset of Governmental Conflict

```
Mean estimation                    Number of obs    =   95718

            0: govonset = 0
            1: govonset = 1

--------------------------------------------------------------
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
Local Elevation
           0 |   640.8905    1.43828      638.0715    643.7095
           1 |    752.843   7.460931      738.2197    767.4664
-------------+------------------------------------------------
Local Elevation Range
           0 |    187.413   .7828656      185.8786    188.9474
```

```
           1 |   190.7436   4.147907      182.6137    198.8734
-------------+----------------------------------------------
Local Elevation Slope as %
           0 |   .5846469   .0031825      .5784093    .5908845
           1 |   .5102022   .0154539      .4799127    .5404917
-------------+----------------------------------------------
Local Elevation Hills (middle values most positive)
           0 |   .7493001   .0004743      .7483706    .7502297
           1 |   .7633151   .0025498      .7583175    .7683127
-------------+----------------------------------------------
Hills Elevation Range
           0 |    .740519   .0004935      .7395518    .7414861
           1 |   .7561137   .0026403      .7509388    .7612886
-------------+----------------------------------------------
Hills Slope as %
           0 |   .7477727   .0004841      .7468239    .7487215
           1 |   .7413131   .0027195      .7359829    .7466433
-------------+----------------------------------------------
Local Forest Density Scale
           0 |   .6373362   .0027319      .6319818    .6426907
           1 |   1.288548    .018409      1.252466    1.324629
-------------+----------------------------------------------
Local Forest Middle values are most positive
           0 |   .5968964   .0003733      .5961647    .5976281
           1 |    .609589   .0019245      .6058171    .6133609
-------------+----------------------------------------------
Local Forest Extreme values are most positive
           0 |   .6728072   .0011895      .6704757    .6751386
           1 |   .6215702   .0061296      .6095562    .6335842
```

Mean Estimates for Local Terrain Data and Onset of Territorial Conflict

Mean estimation Number of obs = 95718

0: teronset = 0

1: teronset = 1

Over	Mean	Std. Err.	[95% Conf.	Interval]
Local Elevation				
0	**647.1202**	**1.43415**	**644.3093**	**649.9312**
1	**494.7237**	**5.762028**	**483.4302**	**506.0172**
Local Elevation Range				
0	**188.8956**	**.7802266**	**187.3664**	**190.4248**
1	**110.6265**	**3.264813**	**104.2275**	**117.0255**
Local Elevation Slope as %				
0	**.5866903**	**.0031659**	**.5804852**	**.5928955**
1	**.3364152**	**.013116**	**.3107081**	**.3621224**
Local Elevation Hills (middle values most positive)				
0	**.748403**	**.0004708**	**.7474802**	**.7493257**
1	**.8246479**	**.0028429**	**.8190759**	**.8302198**

```
Hills Elevation Range
           0 |   .7412117   .0004891       .740253    .7421704
           1 |   .7299341   .0037985      .7224891    .7373791
-------------+------------------------------------------------
Hills Slope as %
           0 |   .7474819   .0004813      .7465385    .7484253
           1 |   .7523735   .0033946      .7457201    .7590269
-------------+------------------------------------------------
Local Forest Density Scale
           0 |   .6645221   .0027713      .6590903    .6699539
           1 |   .2925512   .0108765      .2712334     .313869
-------------+------------------------------------------------
Local Forest Middle values are most positive
           0 |   .5975359     .00037      .5968107    .5982611
           1 |   .5840568   .0027113      .5787427    .5893708
-------------+------------------------------------------------
Local Forest Extreme values are most positive
           0 |    .670477   .0011781      .6681679     .672786
           1 |    .710614   .0090703      .6928364    .7283917
--------------------------------------------------------------
```

Mean Estimates for Statewide Forest Data and Onset of Governmental Conflict

```
Mean estimation                    Number of obs    =   95886

            0: govonset = 0
            1: govonset = 1
```

```
               (Std. Err. adjusted for 302 clusters in c_year)
--------------------------------------------------------------
             |               Robust
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
Local Forest Density Scale
           0 |   .6368264   .0603256      .5181131    .7555396
           1 |   1.275081   .3872707      .5129799    2.037182
--------------------------------------------------------------
```

Mean Estimates for Statewide Forest Data and Onset of Territorial Conflict

```
Mean estimation                    Number of obs    =   95886

            0: teronset = 0
            1: teronset = 1

               (Std. Err. adjusted for 302 clusters in c_year)
--------------------------------------------------------------
             |               Robust
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
Local Forest Density Scale
           0 |   .6636861   .0629741      .5397608    .7876114
           1 |   .2898836    .081636      .1292341    .4505332
--------------------------------------------------------------
```

Relation to Occurrence of Conflict

In general, the findings from the onset of conflict are similar to those from the occurrence of conflict, but there is at least one notable exception. The data for elevation range, which was not significant for government conflict onset, is significant for government conflict occurrence. While this data is not structured appropriately to test the hypothesis that conflicts last longer in areas with greater ranges in elevation, it does appear that conflict has some tendency to persist in territory with greater elevation changes.

Also, the mean estimate for the forest cover in conflict and non-conflict cells, both governmental and territorial are closer for occurrence of conflict than they were for onset of conflict. This distinction suggests that conflicts have a tendency to onset in more extreme conditions but then extend into more mild conditions, which is where they persist. While this finding may appear circular – as a conflict endures, its characteristics become more average as it spreads – the underlying assumption is much more interestingly. Namely, without spreading conflict appears to become more short-lived. This finding cannot be fully tested from this data structure, but it is an interesting suggestion.

(As a side note, one of the methods considered for this thesis but not pursued was cellular automata – the method by which a contagion, be it disease, wildfire, or civil conflict, spreads from one area to another. This path may be fruitful for other researchers seeking to model the relationship between the onset of a conflict and its actual territorial reach over time.)

Mean Estimates for Local Terrain Data and Occurrence of Governmental Conflict

```
Mean estimation                    Number of obs    =   95718

            0: govoccur = 0
            1: govoccur = 1

--------------------------------------------------------------
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
Local Elevation
           0 |    628.465    1.50937      625.5067    631.4234
           1 |   747.5093   3.904629      739.8562    755.1623
-------------+------------------------------------------------
Local Elevation Range
           0 |   185.4834   .8312398      183.8542    187.1126
           1 |   200.6504    2.02294      196.6855    204.6154
-------------+------------------------------------------------
Local Elevation Slope as %
           0 |   .5793657   .0033714      .5727578    .5859735
           1 |   .6011378    .008221      .5850246    .6172509
-------------+------------------------------------------------
Local Elevation Hills
           0 |    .752896   .0005025       .751911    .7538809
           1 |   .7294091   .0012358      .7269871    .7318312
-------------+------------------------------------------------
```

```
Hills Elevation Range
            0 |   .7392522   .0005248      .7382236    .7402809
            1 |   .7523762   .0012625      .7499018    .7548506
-------------+----------------------------------------------------
Hills Slope as %
            0 |   .7471689   .0005126      .7461643    .7481735
            1 |   .7501433   .0012955      .7476042    .7526823
-------------+----------------------------------------------------
Local Forest Density Scale
            0 |   .6385253    .002935      .6327728    .6442779
            1 |   .7834924   .0074191       .768951    .7980338
-------------+----------------------------------------------------
Local Forest Middle
            0 |    .592769   .0003906      .5920036    .5935345
            1 |   .6265258   .0010194      .6245279    .6285237
-------------+----------------------------------------------------
Local Forest Extreme
            0 |   .6864409   .0012463      .6839982    .6888836
            1 |   .5727344   .0032076      .5664475    .5790212
------------------------------------------------------------------
```

Mean Estimates for Local Terrain Data and Occurrence of Governmental Conflict

```
Mean estimation                    Number of obs    =   95718

             0: teroccur = 0
             1: teroccur = 1
```

Over	Mean	Std. Err.	[95% Conf.	Interval]
Local Elevation				
0	**638.8311**	**1.454613**	**635.9801**	**641.6821**
1	**718.4134**	**5.799253**	**707.047**	**729.7799**
Local Elevation Range				
0	186.2029	.7900101	184.6545	187.7513
1	204.8737	3.249666	198.5044	211.243
Local Elevation Slope as %				
0	.5784128	.0031939	.5721528	.5846727
1	.6333945	.0136272	.6066852	.6601037
Local Elevation Hills				
0	**.7474348**	**.0004813**	**.7464915**	**.7483781**
1	**.7802099**	**.0018287**	**.7766258**	**.7837941**
Hills Elevation Range				
0	**.7437091**	**.0005016**	**.742726**	**.7446923**
1	**.7054523**	**.0018491**	**.7018282**	**.7090765**
Hills Elevation Slope				
0	**.7487969**	**.0004936**	**.7478294**	**.7497644**
1	**.7313573**	**.0018192**	**.7277917**	**.7349229**
Local Forest Density Scale				

```
          0 |   .6754981   .0028965      .6698209    .6811752
          1 |   .4268694   .0060643      .4149834    .4387555
------------+-------------------------------------------------
Local Forest Middle
          0 |   .5941257   .0003759       .593389    .5948624
          1 |   .6391538   .0014872      .6362389    .6420687
------------+-------------------------------------------------
Local Forest Extreme
          0 |   .6815752   .0011975       .679228    .6839223
          1 |    .534102   .0047353      .5248208    .5433832
--------------------------------------------------------------
```

Distance to Capital and Borders

In the case of distance to the state capital and international borders, advanced GIS technology has helped generate measures not only of the straight line (or "as the crow flies") distance between grid cells and capitals and borders, but it has also allowed for the creation of improved measures of distance that take into account the effect of different elevations changes in the path between the grid cell and the capital or border. This improvement in the variable specification is more than simply an interaction term – it does not just multiply two variables together – instead it actually determines the least difficult path from the grid cell to the capital or border taking into account elevation changes. For future studies of conflict, it is of interest whether this improved specification makes any difference, and, if so, in which cases.

There are some interaction variables in this section that are also particularly important. As described earlier in the literature review and hypotheses, the interaction terms with border distance help to tease out the three main explanations that underlie most arguments for the importance of international borders – the presence of safe havens, the presence of a common ethnic group, and the presence of resources that can be smuggled.

Correlation Table for Distance to Capital and Borders

(obs=95886)

	capdist	cappat~t	borddist	bordpa~t	ca~ddist	capbor..
capdist	1.0000					
cappathdist	0.9681	1.0000				
borddist	0.0598	0.0613	1.0000			
bordpathdist	0.0155	0.0496	0.9463	1.0000		
capborddist	0.9627	0.9311	-0.2123	-0.2412	1.0000	
capbordpat~t	0.9366	0.9579	-0.2126	-0.2394	0.9745	1.0000

While earlier correlation tables have not had significant correlations, the table above does contain some meaningful information. Most importantly, it shows that the metric designed to measure the location of grid cell relative to both the capital and state borders by subtracting the capital distance from the border distance produced a variable that is highly correlated with capital distance alone. This result indicates that in general capitals are located close to borders, giving the constructed variable a consistent positive correlation with simple distance to the capital. Even though the correction in the capital-border distance difference variable is slight, as is the distinction between simple straight line distance and the cost path distance, it is worth considering what impact those corrections might have on creating models with greater explanatory power.

Relation to Onset of Conflict

The tables below reveal the value of accounting for elevation changes in assessing the distance between grid cells, capitals, and state borders. It is not surprising that the basic straight line distances – what is more relevant is that the path distances appear to be more significant. While it is difficult to measure with total certainty from these tables alone, the path distance measures in general appear to make a stronger distinction between the means for conflict and non-conflict grid cells. The strongest example of this situation is the measure of the capital-border distance difference variable for government conflict, which is not significant as a straight line distance, but when elevation changes are taken in the consideration, the difference becomes significant enough to create distinct means for the variable with 95% confidence. The aim of this thesis is not to cleverly specify variables into significance, but, in this case, the path specification improves the match between the data and the underlying theory.

Mean Estimates for Location Data and Onset of Governmental Conflict

```
Mean estimation                    Number of obs    =   95886

            0: govonset = 0
```

```
            1: govonset = 1


--------------------------------------------------------------
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
capdist      |
           0 |   689549.1   1416.048      686773.6    692324.5
           1 |   723306.4    7299.95      708998.6    737614.3
-------------+------------------------------------------------
cappathdist  |
           0 |   903677.2   1756.754      900233.9    907120.4
           1 |   978007.8   9181.517      960012.1    996003.5
-------------+------------------------------------------------
borddist     |
           0 |   137394.9   390.3937      136629.7    138160.1
           1 |   155432.4   2253.033      151016.4    159848.3
-------------+------------------------------------------------
bordpathdist |
           0 |   187670.4   517.7991      186655.5    188685.2
           1 |   218062.3   3015.941      212151.1    223973.5
-------------+------------------------------------------------
capborddist  |
           0 |   552154.1    1447.03        549318    554990.3
           1 |   567874.1    7380.42      553408.5    582339.6
-------------+------------------------------------------------
capbordpat~t |
           0 |   716006.8   1808.671      712461.8    719551.8
           1 |   759945.5   9231.964        741851    778040.1
--------------------------------------------------------------
```

Another important distinction comes out in this means estimation table for territorial conflict onset and capital/border distances: contrary to hypotheses and general theory, but consistent with previous findings by other studies, distance to a state border does not have a distinguishable mean between territorial conflict and non-conflict cells. The more significant factor in determining the likelihood of conflict is the distance from the capital.

Mean Estimates for Location Data and Onset of Territorial Conflict

```
Mean estimation                    Number of obs    =   95886

            0: teronset = 0
            1: teronset = 1
```

Over	Mean	Std. Err.	[95% Conf.	Interval]
capdist				
0	**685051.3**	**1401.815**	**682303.8**	**687798.9**
1	**1001271**	**7915.616**	**985756.9**	**1016786**
cappathdist				
0	**900336.7**	**1746.005**	**896914.5**	**903758.8**
1	**1224030**	**7836.554**	**1208670**	**1239390**
borddist				
0	137917.1	388.7134	137155.3	138679

```
           1 |   140798.8   2711.473      135484.4    146113.3
-------------+--------------------------------------------------
bordpathdist |
           0 |   188940.4   516.6394      187927.8      189953
           1 |   171648.6   3199.607      165377.4    177919.8
-------------+--------------------------------------------------
capborddist  |
           0 |   547134.2   1430.743      544329.9    549938.4
           1 |   860472.6   8932.108      842965.8    877979.4
-------------+--------------------------------------------------
capbordpat~t |
           0 |   711396.3   1793.644      707880.7    714911.8
           1 |    1052381   9425.942       1033907     1070856
----------------------------------------------------------------
```

Relation to Occurrence of Conflict

The table below on the occurrence of government conflict and capital/border distance contains an interesting distinction. While *onset* of governmental conflict was more likely further from the capital, the *occurrence* of governmental conflict is more likely closer to the capital. This finding would seems to indicate two dynamics: first, while government conflicts may originate in the periphery, they march towards the established state center, and second, the conflicts that endure are those that can proceed towards the state capital – governmental conflicts that stay on the periphery are generally quickly ended, it appears. As with the onset of conflict, the distinction between straight line distance and path distance becomes the difference in the breakpoint of 95% confidence for the capital-border difference distance variable.

Mean Estimates for Location Data and Occurrence of Governmental Conflict

```
Mean estimation                     Number of obs    =   95886

            0: govoccur = 0
            1: govoccur = 1

--------------------------------------------------------------
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
capdist      |
           0 |   692820.5   1527.701      689826.2    695814.7
           1 |     676475   3201.332      670200.5    682749.6
-------------+------------------------------------------------
cappathdist  |
           0 |   908602.9   1890.208      904898.1    912307.7
           1 |   889541.1   4094.368      881516.2    897566.1
-------------+------------------------------------------------
borddist     |
           0 |   139217.1   420.7871      138392.4    140041.9
           1 |   129931.6   921.6708      128125.1    131738.1
-------------+------------------------------------------------
bordpathdist |
           0 |   189419.2   556.6311      188328.2    190510.2
           1 |   183596.1   1259.396      181127.7    186064.5
```

```
-------------+------------------------------------------------
capborddist  |
           0 |   553603.3   1558.216      550549.2    556657.4
           1 |   546543.4    3318.22      540039.7    553047.1
-------------+------------------------------------------------
capbordpat~t |
           0 |   719183.7   1945.615      715370.3      722997
           1 |     705945   4190.116      697732.4    714157.6
--------------------------------------------------------------
```

Mean Estimates for Location Data and Occurrence of Territorial Conflict

```
Mean estimation                    Number of obs    =   95886

            0: teroccur = 0
            1: teroccur = 1

--------------------------------------------------------------
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
capdist      |
           0 |   673045.7   1435.498      670232.1    675859.3
           1 |   922014.1   4731.395      912740.6    931287.6
-------------+------------------------------------------------
cappathdist  |
           0 |   884044.9   1792.092      880532.5    887557.4
           1 |    1195593   5290.414       1185224     1205962
```

borddist				
0	**137577.1**	**400.5963**	**136791.9**	**138362.3**
1	**143113.1**	**1378.068**	**140412.1**	**145814.1**
bordpathdist				
0	**187873.8**	**527.7886**	**186839.3**	**188908.2**
1	**198668.5**	**2009.705**	**194729.5**	**202607.5**
capborddist				
0	**535468.6**	**1458.093**	**532610.7**	**538326.4**
1	**778901.1**	**5319.428**	**768475**	**789327.1**
capbordpat~t				
0	**696171.2**	**1829.357**	**692585.6**	**699756.7**
1	**996924.2**	**6347.002**	**984484.2**	**1009364**

Population

Population statistics are generally quite comparable from the local to the national level, and examining them presents a clear case of the differing impact of local-level and state-level variables. The data below reflect analyses for different conflict and non-conflict means with local and state variable assessed separately so that the state-level variables could have the standard errors clustered, as has occurred in other sections, while local-level variables do not. The results are striking: while aggregate data on population appears to have little impact on anticipating conflict within countries, the local level data has a strong impact.

Moreover, the data on local population growth suggests that the theoretical argument that population changes drive conflict through changing resources allocations and social structures, rather than by making the population more difficult for the government to penetrate, may be more credible based on these outcomes. Below is the correlation table for the variables that are used in this section.

Correlation Table for Local and Statewide Population Data

(obs=92961)

	poplocal	poplnl~l	cow_tpop	cow_upop	popden~e	poprat~l	poprat~e
poplocal	1.0000						
poplnlocal	0.4306	1.0000					
cow_tpop	0.1714	0.2221	1.0000				
cow_upop	0.1053	0.0251	0.8090	1.0000			
popdensest~e	0.3808	0.3227	0.4447	0.2662	1.0000		
popratelocal	0.0332	0.1165	0.0353	-0.0148	0.0366	1.0000	
popratestate	0.0286	0.0967	0.0887	-0.0158	0.0814	0.4476	1.0000

Relation to Onset of Conflict

In the question of whether local or state population factors have a more significant impact, there is considerable room for question. With some compelling theories and evidence at the aggregate and disaggregate levels, it would seem likely that both local and state populations would make a difference. However, when the means were compared between conflict and non-conflict cells, there is strong evidence to suggest that the local impact of population is much more significant. On the question of local population impacts, the evidence that higher local population (effectively population density) is associated with government conflict is not surprising because many governmental challenges originate in major cities, if not the capital. On the other side, territorial conflict lives up to its common conception as being based in more remote areas – the conflict tends to be located in much less densely populated areas. But, the more interesting issues in this data – and one that was not considered in the earlier Buhaug-Rød study because it covered so many years in a single time-frame is that of local population growth. There is evidence that slower population growth in dense areas may be associated with government conflict occurrence, and stronger evidence that higher local population growth is associated with territorial conflict both onset and occurrence. Without local economic data it is difficult to confirm the following theory, but it may be that lower growth rates in cities indicates weaker economic growth, spurring movements for governmental change, and higher growth rates in more remote areas places new strains on resources and upsets existing social dynamic, moving a previously remote area to separate itself from the state.

Mean Estimates for Local Population Data and Onset of Governmental Conflict

```
Mean estimation                    Number of obs    =   95886

            0: govonset = 0
```

```
          1: govonset = 1

--------------------------------------------------------------
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
poplocal     |
           0 |   319.0872   4.451003      310.3632    327.8111
           1 |   420.7383   23.21982      375.2278    466.2489
-------------+------------------------------------------------
popratelocal |
           0 |   .0212039   .0000986      .0210107    .0213971
           1 |   .0219565   .0003045      .0213598    .0225533
--------------------------------------------------------------
```

Mean Estimates for Local Population Data and Onset of Territorial Conflict

```
Mean estimation                    Number of obs    =   95886

          0: teronset = 0
          1: teronset = 1

--------------------------------------------------------------
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
poplocal     |
           0 |   326.6577   4.447773      317.9401    335.3753
           1 |   80.19361   7.424963      65.64076    94.74645
-------------+------------------------------------------------
```

```
poprate local |
          0 |   .0211493   .0000966        .02096    .0213386
          1 |   .0256077   .0008092      .0240217    .0271937
--------------------------------------------------------------
```

Mean Estimates for Statewide Population Data and Onset of Governmental Conflict

```
Mean estimation                     Number of obs    =   92961

            0: govonset = 0
            1: govonset = 1

               (Std. Err. adjusted for 293 clusters in c_year)
--------------------------------------------------------------
             |             Robust
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
cow_tpop     |
           0 |   18410.67   1425.277      15605.56    21215.79
           1 |   25994.58   9769.094      6767.814    45221.34
-------------+------------------------------------------------
cow_upop     |
           0 |   3299.036   318.4876      2672.214    3925.859
           1 |   5577.784   2410.879      832.8812    10322.69
-------------+------------------------------------------------
popdensest~e |
           0 |   326.4728   24.91779      277.4315     375.514
           1 |   303.4792   63.18711      179.1193    427.8391
```

```
-------------+------------------------------------------------
popratestate |
           0 |   .0215219   .0012137      .0191332    .0239107
           1 |   .0195832    .003973      .0117638    .0274026
--------------------------------------------------------------
```

Mean Estimates for Statewide Population Data and Onset of Territorial Conflict

```
Mean estimation                    Number of obs    =   92961

            0: teronset = 0
            1: teronset = 1

               (Std. Err. adjusted for 293 clusters in c_year)
--------------------------------------------------------------
             |               Robust
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
cow_tpop     |
           0 |   18653.18   1445.245      15808.76     21497.6
           1 |   18982.87   5446.237      8264.018    29701.73
-------------+------------------------------------------------
cow_upop     |
           0 |   3400.934   329.6725      2752.099     4049.77
           1 |   1902.127   560.4581      799.0778    3005.177
-------------+------------------------------------------------
popdensest~e |
           0 |   326.4646   24.46376       278.317    374.6122
```

```
           1 |   285.4467   87.79147      112.6625     458.231
-------------+--------------------------------------------------
popratestate |
           0 |    .021461   .0012013      .0190968    .0238252
           1 |   .0213174   .0043176      .0128198    .0298149
----------------------------------------------------------------
```

To check that these results suggesting no impact for population size, which would be surprising, two additional dependent conditions were checked – state-level onset of government conflict and state-level onset of territorial conflict. As the results below indicate, there does not appear to be a state-level effect either. The reason for finding a distinction between this study and previous studies may be the different thresholds at which conflict is measured. In most of the previous studies that highlighted the impact of population, the dependent variable used was civil war – with more than 1000 battle-related deaths per annum – as opposed to civil conflict – any organized conflict with more than 25 battle-related deaths per annum – which this study uses based on the Uppsala/PRIO dataset.

Mean Estimates for State Population Data and Onset of Governmental Conflict Anywhere in the State

```
Mean estimation                    Number of obs    =   92961

            0: govonsetstate = 0
            1: govonsetstate = 1

               (Std. Err. adjusted for 293 clusters in c_year)
```

Over	Mean	Robust Std. Err.	[95% Conf.	Interval]
cow_tpop				
0	18670.82	1565.092	15590.53	21751.11
1	18580.94	3497.403	11697.63	25464.26
cow_upop				
0	3418.097	344.1728	2740.723	4095.471
1	3075.586	984.4884	1137.993	5013.179
popdensest~e				
0	332.6416	27.21359	279.0819	386.2012
1	279.2349	38.75304	202.9642	355.5056
popratestate				
0	.0214624	.0013154	.0188736	.0240512
1	.0214315	.0025272	.0164576	.0264054

Mean Estimates for State Population Data and Onset of Governmental Conflict Anywhere in the State

```
Mean estimation                    Number of obs    =   92961

            0: teronsetstate = 0
            1: teronsetstate = 1
```

```
               (Std. Err. adjusted for 293 clusters in c_year)
--------------------------------------------------------------
             |             Robust
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
cow_tpop     |
           0 |   18506.53    1483.95      15585.93    21427.12
           1 |   21205.27   5641.121      10102.86    32307.68
-------------+------------------------------------------------
cow_upop     |
           0 |   3455.014   341.3376       2783.22    4126.807
           1 |   2017.446   410.1385      1210.243    2824.648
-------------+------------------------------------------------
popdensest~e |
           0 |   326.3209   25.10577      276.9097    375.7321
           1 |   315.6851   95.70985      127.3165    504.0537
-------------+------------------------------------------------
popratestate |
           0 |   .0215163   .0012411      .0190737     .023959
           1 |   .0204917   .0037046      .0132006    .0277828
--------------------------------------------------------------
```

Relation to Occurrence of Conflict

As mentioned earlier, the results on local population growth are compelling and suggest that the conflict dynamics may be impacted as much or more by changes in population as by the actual population numbers themselves.

Mean Estimates for Local Population Data and Occurrence of Governmental Conflict

```
Mean estimation                         Number of obs    =   95886

            0: govoccur = 0
            1: govoccur = 1

--------------------------------------------------------------
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
poplocal     |
           0 |   317.8971   4.515328      309.0471    326.7471
           1 |   350.7453   14.58001      322.1687     379.322
-------------+------------------------------------------------
popratelocal |
           0 |   .0216623   .0001015      .0214633    .0218612
           1 |   .0184329   .0002855      .0178733    .0189925
--------------------------------------------------------------
```

Mean Estimates for Local Population Data and Occurrence of Territorial Conflict

```
Mean estimation                         Number of obs    =   95886

            0: teroccur = 0
            1: teroccur = 1

--------------------------------------------------------------
```

```
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
poplocal     |
           0 |   327.6539   4.641828       318.556    336.7518
           1 |   252.0445   10.05946       232.328    271.7609
-------------+------------------------------------------------
popratelocal |
           0 |   .0210975      .0001      .0209014    .0212936
           1 |   .0229435    .000335      .0222869       .0236
--------------------------------------------------------------
```

Mean Estimates for State Population Data and Occurrence of Governmental Conflict

```
Mean estimation                     Number of obs    =   92961

            0: govoccur = 0
            1: govoccur = 1

               (Std. Err. adjusted for 293 clusters in c_year)
--------------------------------------------------------------
             |             Robust
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
cow_tpop     |
           0 |   19041.74   1530.339      16029.85    22053.64
           1 |   16284.81   3504.134      9388.253    23181.37
-------------+------------------------------------------------
cow_upop     |
```

```
           0 |   3407.305   331.8555      2754.173    4060.437
           1 |   3165.158   923.2988      1347.994    4982.322
-------------+------------------------------------------------
popdensest~e |
           0 |   335.6741   27.33397      281.8776    389.4707
           1 |   263.9371   31.83889      201.2743    326.5999
-------------+------------------------------------------------
popratestate |
           0 |   .0216978   .0012161      .0193044    .0240912
           1 |   .0199725   .0026329      .0147906    .0251544
--------------------------------------------------------------
```

Mean Estimates for State Population Data and Occurrence of Territorial Conflict

```
Mean estimation                    Number of obs    =   92961

            0: teroccur = 0
            1: teroccur = 1

               (Std. Err. adjusted for 293 clusters in c_year)
--------------------------------------------------------------
             |               Robust
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
cow_tpop     |
           0 |   18157.62    1458.21      15287.69    21027.55
           1 |   26313.07    3742.81      18946.76    33679.37
-------------+------------------------------------------------
cow_upop     |
```

```
           0 |   3410.796   340.5364      2740.578    4081.013
           1 |    2807.65   481.6791      1859.647    3755.653
-------------+--------------------------------------------------
popdensest~e |
           0 |   323.3452   24.68021      274.7716    371.9189
           1 |   361.9496   63.60011      236.7769    487.1223
-------------+--------------------------------------------------
popratestate |
           0 |   .0214424   .0012446       .018993    .0238919
           1 |   .0217016   .0019047      .0179529    .0254503
----------------------------------------------------------------
```

As before, the surprising non-result on the impact of aggregate population on conflict is checked again using state-level conflict occurrence measures as dependent variables. And, as before, there is still no relationship when state-level dependent variables are used. This finding is surprising, given the results in earlier literature that found a significant role for population. The different thresholds for conflict may be the reason for the non-result. This study uses a threshold of 25 battle-related deaths per annum from the Uppsala/PRIO conflict dataset, but most other population studies have used a dataset of civil wars – those conflicts with more than 1,000 battle-related deaths per annum.

Mean Estimates for State Population Data and Occurrence of Governmental Conflict Anywhere in the State

```
Mean estimation                    Number of obs    =   92961

            0: govoccurstate = 0
            1: govoccurstate = 1
```

```
               (Std. Err. adjusted for 293 clusters in c_year)
--------------------------------------------------------------
             |               Robust
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
cow_tpop     |
           0 |   17644.33   1705.491      14287.72    21000.95
           1 |   21068.79   2605.002      15941.83    26195.75
-------------+------------------------------------------------
cow_upop     |
           0 |   3312.475   349.1492      2625.307    3999.643
           1 |   3519.082   717.6036      2106.751    4931.413
-------------+------------------------------------------------
popdensest~e |
           0 |   325.1958   31.11821      263.9513    386.4402
           1 |   326.9626   35.46546      257.1622    396.7629
-------------+------------------------------------------------
popratestate |
           0 |   .0210821   .0013208      .0184826    .0236815
           1 |   .0223519   .0024847      .0174617    .0272422
--------------------------------------------------------------
```

Mean Estimates for State Population Data and Occurrence of Territorial Conflict Anywhere in the State

```
Mean estimation                    Number of obs    =   92961
```

```
            0: teroccurstate = 0
            1: teroccurstate = 1

               (Std. Err. adjusted for 293 clusters in c_year)
--------------------------------------------------------------
             |               Robust
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
cow_tpop     |
           0 |   17138.62   1679.472      13833.21    20444.02
           1 |   24256.77   2473.804      19388.02    29125.51
-------------+------------------------------------------------
cow_upop     |
           0 |   3355.036   391.5102      2584.497    4125.576
           1 |   3442.419   494.8484      2468.498    4416.341
-------------+------------------------------------------------
popdensest~e |
           0 |   325.9896   28.42524      270.0453    381.9339
           1 |   324.7244   44.51651      237.1105    412.3383
-------------+------------------------------------------------
popratestate |
           0 |     .02186   .0014446      .0190169    .0247031
           1 |   .0199799    .001597      .0168367    .0231231
--------------------------------------------------------------
```

Natural Resources

The data on natural resources comes in three broad categories: aggregate data on the amount of the economy relying on primary commodity exports, grid cell data on the density of petroleum deposits, and grid cell data on the density of non-fuel mineral deposits. As with other categories, this data makes it possible to distinguish between the difference between state-level effects that can occur because of the overall composition of the economy and local-level effects that can occur because of proximity to resources. While many more minerals will be assessed than just diamonds, the table below includes only the general measure for density of diamonds and the measure for the density of secondary diamonds – those most often thought to be lootable and useful for rebels. It is good to see in the table below that the aggregate and disaggregate measures align well: those economies with a high proportion of oil also have low distance to petroleum (a negative correlation) and high petroleum density (a positive correlation). It is also interesting to note that mineral density – both general minerals and secondary diamonds – do not appear to be particularly correlated with petroleum density.

Correlation Table for Local and Statewide Natural Resource Data

(obs=87402)

	gg_sxp	gg_sxp2	gg_oil~p	gg_oil~2	petrod~t	petrod~e	minera~e	diasec~y
gg_sxp	1.0000							
gg_sxp2	0.9652	1.0000						
gg_oilsxp	0.6357	0.6319	1.0000					
gg_oilsxp2	0.6707	0.7032	0.9549	1.0000				
petrodist	0.0245	0.0169	-0.3065	-0.2139	1.0000			
petrodense	0.0553	0.0291	0.2479	0.1414	-0.3911	1.0000		
mineralsde~e	-0.0674	-0.0704	-0.0839	-0.0807	0.0141	0.0018	1.0000	

```
diasecdens~y |   0.1258   0.1213   0.0316   0.0565   0.1180  -0.0538   0.1127
1.0000
```

Petroleum

At the disaggregate level, the data on petroleum shows a few notable trends – and a few non-results that are useful indicators. The results for territorial conflict onset and government conflict occurrence both present an interesting conundrum: conflict appears to be more likely *both* in areas that are closer to petroleum and that have smaller densities of petroleum. On first glance, this situation would seem extraordinarily counterintuitive, but there is a possible reasonable explanation: it may be that conflict is more likely around low concentrations of oil because they do not garner the level of government security and presence that heavy oil installations create. So, in a small concentration of oil installations, the prize is still desirable target for rebels but it also presents an attainable one because government presence is not boosted very high by a large concentration of oil.

Mean Estimates for Local Petroleum Data and Onset of Governmental Conflict

```
Mean estimation                          Number of obs    =   95886


            0: govonset = 0
            1: govonset = 1


--------------------------------------------------------------
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
```

```
petrodist    |
           0 |   633506.5   1405.787      630751.2    636261.8
           1 |   645233.9   7809.426      629927.5    660540.3
-------------+------------------------------------------------
petrodense   |
           0 |   .1555136   .0017588      .1520664    .1589608
           1 |     .15051    .009029      .1328132    .1682067
--------------------------------------------------------------
```

Mean Estimates for Local Petroleum Data and Onset of Territorial Conflict

```
Mean estimation                     Number of obs    =   95886

            0: teronset = 0
            1: teronset = 1

--------------------------------------------------------------
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
petrodist    |
           0 |   635707.1   1400.272      632962.6    638451.6
           1 |   531927.8   8031.596      516185.9    547669.6
-------------+------------------------------------------------
petrodense   |
           0 |   .1571454   .0017561      .1537034    .1605874
           1 |   .0554821   .0035754      .0484743      .06249
--------------------------------------------------------------
```

Mean Estimates for Local Petroleum Data and Occurrence of Governmental Conflict

```
Mean estimation                    Number of obs    =   95886

            0: govoccur = 0
            1: govoccur = 1

--------------------------------------------------------------
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
petrodist    |
           0 |   639401.8   1495.413      636470.8    642332.8
           1 |   598351.3   3620.017      591256.1    605446.5
-------------+------------------------------------------------
petrodense   |
           0 |    .164745   .0019305      .1609612    .1685287
           1 |   .0949465   .0031945      .0886854    .1012076
--------------------------------------------------------------
```

Mean Estimates for Local Petroleum Data and Occurrence of Territorial Conflict

```
Mean estimation                    Number of obs    =   95886

            0: teroccur = 0
            1: teroccur = 1
```

```
-------------------------------------------------------------
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+-----------------------------------------------
petrodist    |
           0 |   633997.5   1439.059      631176.9      636818
           1 |     632320   5016.891        622487    642153.1
-------------+-----------------------------------------------
petrodense   |
           0 |   .1620644   .0018405      .1584571    .1656717
           1 |    .067018   .0031664       .060812     .073224
-------------------------------------------------------------
```

Non-Fuel Minerals

The data on non-fuel minerals includes a wider array, so a new correlation table has been produced below. It indicates that, unsurprisingly, most mineral deposits are correlated with each other. The one notable exception to this rule is coltan, which has negative correlations with manganese, gold, platinum, iron, and secondary diamonds.

Correlation Table for Non-Fuel Mineral Resource Density

```
(obs=95886)

             | minera~e mangan~e goldde~e copper~e platin~e chromi~e coltan~e ironde~e
diasec~y
-------------+------------------------------------------------------------------------
---------
mineralsde~e |   1.0000
manganesed~e |   0.3090   1.0000
```

```
   golddense |   0.8295   0.1792   1.0000
 copperdense |   0.4590   0.0758   0.2012   1.0000
platinumde~e |   0.7277   0.1382   0.5704   0.2819   1.0000
chromiumde~e |   0.7824   0.1123   0.5374   0.3162   0.7163   1.0000
 coltandense |   0.0663  -0.0277  -0.0050   0.0191  -0.0197   0.0117   1.0000
   irondense |   0.4360   0.3105   0.2111   0.1048   0.3323   0.2628  -0.0381   1.0000
diasecdens~y |   0.1149   0.0918   0.1123   0.0020   0.0613   0.0389  -0.0240   0.4127
1.0000
```

The initial results on various mineral densities and their relation to conflict indicates broadly that areas of governmental conflict onset had lower mineral density, with the notable exception of secondary diamonds and copper. Similarly, areas of territorial conflict onset had lower mineral density with the exception of coltan. That secondary diamonds were associated with *lower* territorial conflict rather than higher conflict is striking, but the result may be driven heavily by data from Sierra Leone which is classified as a governmental conflict. These findings broadly suggest that the concern about minerals in general may be generally limited to diamonds, which have much higher value per weight, as discussed earlier. And, although the list of cases of rebels using natural resources for financing is not short, there are many countervailing examples where resource deposits have brought government presence and development, reducing the risk of conflict.

Mean Estimates for Local Non-Fuel Mineral Density and Onset of Governmental Conflict

```
Mean estimation                    Number of obs    =   95886

              0: govonset = 0
```

```
          1: govonset = 1

--------------------------------------------------------------
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
mineralsde~e |
           0 |   .7429204   .0072037      .7288012    .7570397
           1 |   .4586026   .0175219      .4242598    .4929453
-------------+------------------------------------------------
manganesed~e |
           0 |   .0301168   .0005437      .0290512    .0311824
           1 |     .00786   .0007141      .0064604    .0092596
-------------+------------------------------------------------
golddense    |
           0 |   .1593184   .0036498      .1521647     .166472
           1 |   .0238717   .0020544      .0198451    .0278983
-------------+------------------------------------------------
copperdense  |
           0 |   .0807827   .0011083      .0786106    .0829549
           1 |    .116759   .0094761      .0981859     .135332
-------------+------------------------------------------------
platinumde~e |
           0 |   .0271741   .0006967      .0258087    .0285396
           1 |   .0022692   .0006461      .0010028    .0035356
-------------+------------------------------------------------
chromiumde~e |
           0 |   .0603925   .0013294      .0577868    .0629982
           1 |   .0009959   .0001776      .0006478     .001344
-------------+------------------------------------------------
```

```
coltandense  |
           0 |   .0072786   .0001142      .0070548    .0075023
           1 |   .0171453   .0010328      .0151211    .0191695
-------------+------------------------------------------------
irondense    |
           0 |    .036543   .0004284      .0357034    .0373826
           1 |   .0410712   .0027711      .0356399    .0465024
-------------+------------------------------------------------
diasecdens~y |
           0 |   .1010557   .0014228       .098267    .1038444
           1 |   .1910438   .0095027      .1724186    .2096689
--------------------------------------------------------------
```

Mean Estimates for Local Non-Fuel Mineral Density and Onset of Territorial Conflict

```
Mean estimation                     Number of obs    =   95886

            0: teronset = 0
            1: teronset = 1

--------------------------------------------------------------
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
mineralsde~e |
           0 |   .7452104   .0071185      .7312583    .7591626
           1 |   .1024543   .0046729      .0932955    .1116131
-------------+------------------------------------------------
```

```
manganesed~e |
           0 |   .0299111   .0005363      .0288599    .0309623
           1 |   .0014505   .0002714      .0009186    .0019824
-------------+-----------------------------------------------
golddense    |
           0 |   .1575988   .0035981      .1505465    .1646511
           1 |   .0109529   .0007258      .0095303    .0123755
-------------+-----------------------------------------------
copperdense  |
           0 |   .0833454   .0011341      .0811226    .0855681
           1 |   .0027428   .0004007      .0019575    .0035281
-------------+-----------------------------------------------
platinumde~e |
           0 |    .026852    .000687      .0255055    .0281984
           1 |   .0002276   .0000694      .0000916    .0003636
-------------+-----------------------------------------------
chromiumde~e |
           0 |   .0593005   .0013101      .0567328    .0618683
           1 |   .0141726   .0020354      .0101832    .0181619
-------------+-----------------------------------------------
coltandense  |
           0 |   .0075387   .0001152      .0073129    .0077646
           1 |   .0105621   .0012835      .0080465    .0130776
-------------+-----------------------------------------------
irondense    |
           0 |   .0371554   .0004307      .0363111    .0379996
           1 |    .010558   .0013891      .0078355    .0132805
-------------+-----------------------------------------------
diasecdens~y |
```

```
        0 |   .1057679   .0014358      .1029538    .1085821
        1 |    .000537   .0001366      .0002693    .0008048
--------------------------------------------------------------
```

On the issue of the occurrence of conflict, there is a marked shift in results from onset in the case of governmental conflict that does not occur for territorial conflict. In the case of governmental conflict onset, lower mineral levels were generally associated with conflict, but in the case of governmental conflict occurrence, higher mineral levels appear to be slightly more associated with conflict. This suggests that that the "prize" element of governmental conflict may be at play – the tendency for rebels to focus on capturing the key "prizes" or strategic targets in a country. It may also suggest that governmental conflicts with higher mineral concentrations tend to last longer – an idea that has been suggested by other researchers (Fearon 2004)((Fearon 2004)Fearon 2004). The other interesting aspect of these results is that territorial conflict occurrence continues to be less likely in areas with mineral resources, as with territorial conflict onset. The reason behind this continuity, which is not shared with governmental conflict, may be that territorial conflicts are less interested in the strategic targets within a country in general and they have less interest in spreading through the entirety of the country – in short, they are less interested in "prizes." Instead, territorial conflicts tend to remain in the same areas with less attention to mineral deposits.

Mean Estimates for Local Non-Fuel Mineral Density and Occurrence of Governmental Conflict

```
Mean estimation                    Number of obs    =   95886
```

```
        0: govoccur = 0
        1: govoccur = 1

--------------------------------------------------------------
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
mineralsde~e |
           0 |   .7154088   .0071507      .7013936     .729424
           1 |   .8527689   .0242725      .8051951    .9003427
-------------+------------------------------------------------
manganesed~e |
           0 |    .028998   .0005297      .0279598    .0300361
           1 |   .0320591   .0019331      .0282702     .035848
-------------+------------------------------------------------
golddense    |
           0 |   .1522165   .0037365      .1448929      .15954
           1 |    .173023   .0106175      .1522128    .1938332
-------------+------------------------------------------------
copperdense  |
           0 |   .0803549   .0012047      .0779937    .0827162
           1 |   .0920297   .0029279      .0862909    .0977684
-------------+------------------------------------------------
platinumde~e |
           0 |   .0237305   .0006725      .0224124    .0250487
           1 |   .0434463   .0025355      .0384766    .0484159
-------------+------------------------------------------------
chromiumde~e |
           0 |   .0515896   .0012553      .0491292    .0540499
```

```
           1 |   .1029969   .0051209      .0929599    .1130338
-------------+------------------------------------------------
coltandense  |
           0 |   .0072345   .0001226      .0069942    .0074748
           1 |   .0098918   .0003377        .00923    .0105536
-------------+------------------------------------------------
irondense    |
           0 |   .0343743   .0004336      .0335243    .0352242
           1 |    .051564   .0014599      .0487026    .0544254
-------------+------------------------------------------------
diasecdens~y |
           0 |   .0926947    .001453       .089847    .0955425
           1 |   .1760904   .0047122      .1668544    .1853263
--------------------------------------------------------------
```

Mean Estimates for Local Non-Fuel Mineral Density and Occurrence of Territorial Conflict

```
Mean estimation                    Number of obs    =   95886

            0: teroccur = 0
            1: teroccur = 1

--------------------------------------------------------------
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
mineralsde~e |
           0 |   .7648039   .0075024      .7500993    .7795085
```

```
           1 |    .326873   .0068549      .3134375    .3403084
-------------+----------------------------------------------
manganesed~e |
           0 |   .0313675   .0005663      .0302576    .0324774
           1 |   .0036361   .0002229      .0031993     .004073
-------------+----------------------------------------------
golddense    |
           0 |   .1648895   .0038014      .1574389    .1723401
           1 |   .0250216   .0007897      .0234737    .0265694
-------------+----------------------------------------------
copperdense  |
           0 |   .0833826   .0011833      .0810633     .085702
           1 |   .0627426   .0025574        .05773    .0677551
-------------+----------------------------------------------
platinumde~e |
           0 |   .0282457   .0007258      .0268232    .0296682
           1 |   .0018598   .0001619      .0015425    .0021771
-------------+----------------------------------------------
chromiumde~e |
           0 |   .0615507   .0013822      .0588415    .0642599
           1 |    .018416   .0011365      .0161885    .0206435
-------------+----------------------------------------------
coltandense  |
           0 |   .0077628   .0001192      .0075291    .0079964
           1 |   .0053437   .0004584      .0044453    .0062422
-------------+----------------------------------------------
irondense    |
           0 |    .038711   .0004532      .0378227    .0395993
           1 |   .0100384   .0005982       .008866    .0112108
```

```
-------------+------------------------------------------------
diasecdens~y |
           0 |   .1116462   .0015153      .1086763    .1146161
           1 |   .0021206   .0001467      .0018331    .0024082
--------------------------------------------------------------
```

Diamonds in Focus

Given the strong findings that secondary diamond densities had in the initial tests, it is worth exploring whether the impact of diamonds is the same for all kinds of diamonds, or, as hypothesized, secondary, alluvial, or more lootable diamonds promote conflict while primary, kimberlite, or less lootable diamonds prevent conflict. In examining this correlation table it is worth noting two points: first, there does not appear to be strong correlations between the densities of lootable and non-lootable diamonds; and second, there is a much stronger correlation between the densities of non-lootable diamonds in general and those with known production than there is with lootable diamonds and those this known production. This distinction suggests that non-lootable diamonds are generally only found when there is reason for production – which is logical because the reason the diamonds are non-lootable is that they are difficult to find and extract from the earth. On the other hand, non-lootable diamonds are less often exploited by large companies or governments, reducing the correlation between non-lootable diamond occurrences and those with known production.

Correlation Table for Political Regime Scores

```
(obs=95886)
```

```
             | dialde~e dialyp~e dianld~e dianly~e diasec~t diasec~y
-------------+------------------------------------------------------
   dialdense |   1.0000
 dialypdense |   0.1495   1.0000
  dianldense |   0.1495   1.0000   1.0000
dianlypdense |   0.0339   0.9919   0.9919   1.0000
  diasecdist |  -0.1479  -0.1064  -0.1064  -0.0920   1.0000
diasecdens~y |   0.9286   0.2220   0.2220   0.1233  -0.2721   1.0000
```

Relation to Onset of Conflict

The relationship between conflict onset and diamonds is strong – but not all diamond deposits are created equal when it comes to their effect on conflict. While the sets of all lootable diamonds and all secondary diamonds (which largely overlap) have strong positive relationship with onset of governmental conflict, all other forms of diamond deposits (including "lootable" diamond deposits with known production) have a negative relationship with conflict. Meanwhile, the effect with territorial conflict is more uniform – all territorial conflict onsets were associated with an absence of diamond deposits.

Mean Estimates for Local Diamond Density and Onset of Governmental Conflict

```
Mean estimation                     Number of obs    =   95886

            0: govonset = 0
            1: govonset = 1
```

Over	Mean	Std. Err.	[95% Conf.	Interval]
dialdense				
0	**.908847**	**.0243302**	**.8611601**	**.9565339**
1	**1.486034**	**.1372985**	**1.216931**	**1.755138**
dialypdense				
0	**.2628472**	**.0090515**	**.2451063**	**.280588**
1	**.0741022**	**.0072574**	**.0598778**	**.0883267**
dianldense				
0	**.2628472**	**.0090515**	**.2451063**	**.280588**
1	**.0741022**	**.0072574**	**.0598778**	**.0883267**
dianlypdense				
0	**.1904697**	**.0075451**	**.1756813**	**.2052581**
1	**.0304343**	**.0031909**	**.0241803**	**.0366884**
diasecdist				
0	**658529.9**	**1669.161**	**655258.3**	**661801.4**
1	**380759.5**	**6949.238**	**367139.1**	**394380**
diasecdens~y				
0	**.1010557**	**.0014228**	**.098267**	**.1038444**
1	**.1910438**	**.0095027**	**.1724186**	**.2096689**

Mean Estimates for Local Diamond Density and Onset of Territorial Conflict

```
Mean estimation                    Number of obs    =   95886

            0: teronset = 0
            1: teronset = 1

--------------------------------------------------------------
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
dialdense    |
           0 |   .9438067    .024386      .8960105    .9916029
           1 |   .0000336   .0000219     -9.31e-06    .0000765
-------------+------------------------------------------------
dialypdense  |
           0 |   .2614568   .0089241      .2439656     .278948
           1 |          0          0             .           .
-------------+------------------------------------------------
dianldense   |
           0 |   .2614568   .0089241      .2439656     .278948
           1 |          0          0             .           .
-------------+------------------------------------------------
dianlypdense |
           0 |   .1887099   .0074373      .1741329    .2032869
           1 |          0          0             .           .
-------------+------------------------------------------------
diasecdist   |
           0 |   644352.1   1654.646        641109    647595.2
```

```
           1 |   948296.1   9233.947      930197.7    966394.5
-------------+------------------------------------------------
diasecdens~y |
           0 |   .1057679   .0014358      .1029538    .1085821
           1 |    .000537   .0001366      .0002693    .0008048
--------------------------------------------------------------
```

Relation to Occurrence of Conflict

Occurrence of conflict appears to have slightly different relationship with diamonds than onset. While conflict appeared unlikely to onset near diamonds, there is consistent evidence that governmental conflict is more likely to occur in areas with conflict. This provides more evidence for the “prize” theory of governmental conflict – that while it is unlikely to onset near resources, it will spread towards them as they are strategic prizes in taking over the country and financing a rebellion. The evidence also remains consistent that territorial conflicts do not spread toward resources in the same why because there is less interest in capturing the strategic elements of a country.

Mean Estimates for Local Diamond Density and Occurrence of Governmental Conflict

```
Mean estimation                     Number of obs    =   95886

            0: govoccur = 0
            1: govoccur = 1

--------------------------------------------------------------
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
```

```
dialdense    |
           0 |   .8591731   .0253557      .8094762    .9088699
           1 |   1.364688   .0713758      1.224793    1.504584
-------------+---------------------------------------------------
dialypdense  |
           0 |   .2412815   .0091266      .2233935    .2591696
           1 |    .357012   .0282703      .3016025    .4124216
-------------+---------------------------------------------------
dianldense   |
           0 |   .2412815   .0091266      .2233935    .2591696
           1 |    .357012   .0282703      .3016025    .4124216
-------------+---------------------------------------------------
dianlypdense |
           0 |   .1731806   .0075906      .1583032    .1880581
           1 |   .2639025   .0237647       .217324     .310481
-------------+---------------------------------------------------
diasecdist   |
           0 |   673709.1    1730.79      670316.8    677101.5
           1 |   495292.9   4718.188      486045.3    504540.5
-------------+---------------------------------------------------
diasecdens~y |
           0 |   .0926947    .001453       .089847    .0955425
           1 |   .1760904   .0047122      .1668544    .1853263
-----------------------------------------------------------------
```

Mean Estimates for Local Diamond Density and Occurrence of Territorial Conflict

```
Mean estimation                    Number of obs    =   95886
```

```
            0: teroccur = 0

            1: teroccur = 1


--------------------------------------------------------------

        Over |       Mean   Std. Err.     [95% Conf. Interval]

-------------+------------------------------------------------

dialdense    |

           0 |   .9976046   .0257644      .9471067    1.048103

           1 |   .0000419   .0000122       .000018    .0000659

-------------+------------------------------------------------

dialypdense  |

           0 |   .2763608   .0094304      .2578774    .2948442

           1 |          0          0             .           .

-------------+------------------------------------------------

dianldense   |

           0 |   .2763608   .0094304      .2578774    .2948442

           1 |          0          0             .           .

-------------+------------------------------------------------

dianlypdense |

           0 |   .1994671   .0078597      .1840622     .214872

           1 |          0          0             .           .

-------------+------------------------------------------------

diasecdist   |

           0 |   629941.2   1694.947      626619.2    633263.3

           1 |   909917.6   5489.928      899157.4    920677.8

-------------+------------------------------------------------

diasecdens~y |

           0 |   .1116462   .0015153      .1086763    .1146161
```

```
           1 |   .0021206   .0001467      .0018331    .0024082
--------------------------------------------------------------
```

Form of Government

Measures on the form of government are all collected at the state level, so the tables and analysis presented all have their errors clustered at the country-year level. Because the Polity IV dataset essentially uses categorical variables to then construct scales for democracy and autocracy, the analysis sometimes uses the score solely at the first year of the five-year period to construct a categorical analysis when handling sub-component variables. While it might be tempting to treat these sub-component variables as continuous scales as well, it is clear from the way that they have been coded (as categorical variables) that they cannot reasonably be treated as such.

Correlation Table for Political Regime Scores

```
(obs=91197)

             | p4_de~5y p4_dem~x p4_au~5y p4_aut~x p4_pol..
-------------+---------------------------------------------
 p4_democ_5y |   1.0000
p4_democ_t~x |   0.8300   1.0000
 p4_autoc_5y |  -0.8597  -0.7094   1.0000
p4_autoc_t~x |  -0.7318  -0.8598   0.8750   1.0000
p4_polity2~s |   0.9500   0.8036  -0.9470  -0.8307   1.0000
```

The correlation table above has the expected negative correlations, essentially indicating that countries with more democratic traits have fewer autocratic traits. The interesting discrepancy, though, is that the five-year averages are not more heavily correlated with the mixed average and lag variable generated to give the average of the values prior to the onset of conflict. This situation suggests that there may be a substantial number of cases when including post-conflict political regime scores has affected the overall results.

Relation to Onset of Conflict

The findings for onset of conflict highlight two points that were interesting from the beginning of the model specification. First, in the case of territorial conflict onset, the specification of a mix of averaging and lagging variables so that only pre-conflict scores are included appears to have helped improve the specification of autocratic regime scores from being non-significant to being significant (with different means not sharing 95% confidence interval). Second, the governmental and territorial onset results indicate the value of looking at the subcomponents of the composite scores. While the Polity IV score overall is not significant for either one, but the democratic score is significant for government conflict onset and the autocratic score is significant for territorial conflict onset. This distinction – with democratic scores only significant for governmental conflict and autocratic scores only significant for territorial conflict may also speak to how different government types create conflicts. In other words, in countries where people are excluded from taking part in the government – those with low democracy scores – there is a greater association with attempting to take over the government itself. On the other hand, those governments with little role for civil rights and minority protection – those with high autocracy scores – there is a greater association with attempting to leave the government entirely.

Mean Estimates for Political Regime Scores and Onset of Governmental Conflict

```
Mean estimation                    Number of obs    =   91197

            0: govonset = 0
            1: govonset = 1

               (Std. Err. adjusted for 277 clusters in c_year)
--------------------------------------------------------------
             |              Robust
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
p4_democ_5y  |
           0 |   1.318594   .2312806      .8632954    1.773892
           1 |   .3404095   .2049675     -.0630889    .7439078
-------------+------------------------------------------------
p4_democ_t~x |
           0 |   1.207842   .2360134      .7432268    1.672457
           1 |   .1976293   .1073994     -.0137968    .4090554
-------------+------------------------------------------------
p4_autoc_5y  |
           0 |   6.017095   .2151219      5.593607    6.440583
           1 |   5.995959    .454978      5.100291    6.891627
-------------+------------------------------------------------
```

```
p4_autoc_t~x |
           0 |   6.148725   .2144789      5.726502    6.570947
           1 |   6.514009   .3436872      5.837427     7.19059
-------------+------------------------------------------------
p4_polity2~s |
           0 |  -4.518767   .4227524     -5.350996   -3.686538
           1 |  -5.669935   .5906318     -6.832651    -4.50722
--------------------------------------------------------------
```

Mean Estimates for Political Regime Scores and Onset of Territorial Conflict

```
Mean estimation                     Number of obs    =   91197

            0: teronset = 0
            1: teronset = 1

               (Std. Err. adjusted for 277 clusters in c_year)
--------------------------------------------------------------
             |               Robust
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
p4_democ_5y  |
           0 |   1.268366   .2268287      .8218315      1.7149
           1 |   3.089921   1.374469      .3841471    5.795695
-------------+------------------------------------------------
p4_democ_t~x |
           0 |    1.20738   .2350819      .7445982    1.670161
```

```
           1 |          0          0             .           .
-------------+------------------------------------------------
p4_autoc_5y  |
           0 |   6.046647   .2102744      5.632701    6.460592
           1 |   4.245422   1.333736      1.619834    6.871009
-------------+------------------------------------------------
p4_autoc_t~x |
           0 |   6.139502   .2138941      5.718431    6.560573
           1 |   7.140184   .1437617      6.857176    7.423193
-------------+------------------------------------------------
p4_polity2~s |
           0 |  -4.593363    .413943      -5.40825   -3.778476
           1 |  -1.519368   2.535026     -6.509811    3.471076
--------------------------------------------------------------
```

Relation to Occurrence of Conflict

The only real result from these tables on occurrence of conflict is that not a single score returned a significant result. This result is striking and suggests that there may not be a strong relationship between lasting civil conflict and political regime type. One of the reasons for this confused relationship is that once a conflict begins, scoring for its political regime type becomes much more complicated and often simply stops until a new regime emerges. As a result, the data available on regime type mid-conflict is almost non-existent. This difference may explain why it is possible to get significant results for conflict onset but not for conflict occurrence.

Mean Estimates for Political Regime Scores and Occurrence of Governmental Conflict

```
Mean estimation                     Number of obs    =   91197

            0: govoccur = 0
            1: govoccur = 1

               (Std. Err. adjusted for 277 clusters in c_year)
--------------------------------------------------------------
             |             Robust
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
p4_democ_5y  |
           0 |   1.299404   .2310444      .8445707    1.754237
           1 |   1.292836    .546538      .2169229    2.368748
-------------+------------------------------------------------
p4_democ_t~x |
           0 |   1.171889    .236226      .7068551    1.636923
           1 |   1.312753   .5469192      .2360903    2.389416
-------------+------------------------------------------------
p4_autoc_5y  |
           0 |   6.034477   .2252956      5.590961    6.477994
           1 |   5.871478   .3766782      5.129951    6.613005
-------------+------------------------------------------------
p4_autoc_t~x |
           0 |   6.185641   .2236507      5.745363    6.625919
           1 |   5.915854   .3799347      5.167916    6.663792
-------------+------------------------------------------------
p4_polity2~s |
           0 |  -4.616964   .4333335     -5.470023   -3.763905
           1 |  -3.932771   .8625985      -5.63088   -2.234663
```

```
------------------------------------------------------------------
```

Mean Estimates for Political Regime Scores and Occurrence of Territorial Conflict

```
Mean estimation                    Number of obs    =   91197

            0: teroccur = 0
            1: teroccur = 1

               (Std. Err. adjusted for 277 clusters in c_year)
--------------------------------------------------------------
             |               Robust
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
p4_democ_5y  |
           0 |   1.208653   .2083937      .7984102    1.618896
           1 |   2.486784    .693364       1.12183    3.851738
-------------+------------------------------------------------
p4_democ_t~x |
           0 |   1.133875   .2108786      .7187403     1.54901
           1 |   1.892061   .7067757       .500705    3.283417
-------------+------------------------------------------------
p4_autoc_5y  |
           0 |   6.072704   .2089833        5.6613    6.484108
           1 |   5.277156   .5541355      4.186287    6.368025
-------------+------------------------------------------------
p4_autoc_t~x |
```

```
           0 |   6.174938    .208241      5.764995     6.58488
           1 |   5.908344   .5123403      4.899752    6.916935
-------------+--------------------------------------------------
p4_polity2~s |
           0 |  -4.688026   .3966908      -5.46895   -3.907102
           1 |  -2.617777   1.148604     -4.878914   -.3566397
----------------------------------------------------------------
```

Ethnicity

Governmental conflict appears to onset closer to ethnic borders, which seems reasonable given concerns that ethnic divisions can serve as flashpoints for conflict. In the broader debate about whether ethnicity is relevant to civil conflict at all, this tentative difference would suggest that it has some relevance. Moreover, given that there are sometimes difficult-to-justify ethnic borders that closely mirror state borders, the finding that governmental conflict is more likely to start near an ethnic border (but not near a state border) strengthens the potential finding around ethnic borders. Somewhat surprisingly, territorial conflict appears more likely to onset further from ethnic borders. At first glance, this finding might appear to indicate that ethnicity is less significant to territorial conflicts, but, based on a limited review of conflict histories, it is also possible that territorial conflict onsets in areas where one ethnic group is dominant enough that they develop an identity separate from the state and accumulate the resources to challenge the state. This might suggest that territorial conflicts are more likely in ethnic heartlands rather than divided territory. The findings for occurrence – as opposed to onset – are similar for both governmental and territorial conflict: in both cases, conflict is more likely to occur near an ethnic border. This finding suggests that conflict gravitates towards these dividing lines between groups as the battle is joined.

The findings for conflict occurrence and ethnic border density are more complicated and less clear in their theoretical significance. While governmental conflict appears more likely to occur in areas with high border density, territorial conflict appears less likely to occur in areas of high border density. The finding on territorial conflict holds with the general view that territorial conflicts are more likely to occur in ethnic heartlands, where a consolidated ethnic group – with few other groups to create heterogeneity – has the strength and will to challenge the state. The finding for governmental conflict is less clear. One possible reading is that ethnic border density reflects that a number of smaller consolidated ethnic groups are present, each one of whom seeks to capture the state rather than just their own territory. From a "prize" perspective, this argument seems reasonable: the smaller territory, possibly devoid of natural resources, owned by a small ethnic group is not nearly as attractive a prize nor as sustainable a holding as the entire state, particularly if the only competitors to control of the state are other smaller ethnic groups.

This overall view of the distinction between territorial and governmental conflict holds through the analysis of ethnic group configuration percentage of territory as well. While governmental conflict appears more likely to onset in areas where the ethnic group configuration has approximately 20% of the territory, territorial conflict is more likely when the ethnic group configuration has approximately 24% of the territory. While this distinction may not seem large – both conflict means for ethnic group configuration percentage are well below the no-conflict means – it does indicate that smaller ethnic groups are more likely to seek control over the whole state, while larger ethnic groups are more likely seek their own state outright. Territorial conflict occurrence also does not appear to extend significantly into the territory of majority ethnic group configurations, while the governmental conflicts appear to occur (once onset) in an indistinguishable range of ethnic group configurations.

These findings on ethnic group configuration percentage suggest that ethnicity is in fact an important factor in determining both the likelihood of conflict and potentially the distinction between territorial and governmental conflict. While it remains to be seen whether these findings hold when controlling for other variables, the metrics generated through refined specification of the GREG dataset open up new avenues for future research as well.

Mean Estimates for Ethnic Border Distance and Onset of Governmental Conflict

```
Mean estimation                         Number of obs    =   95886

            0: govonset = 0
            1: govonset = 1

                (Std. Err. adjusted for 15981 clusters in x_y)
--------------------------------------------------------------
             |                Robust
        Over |       Mean   Std. Err.     [95% Conf. Interval]
```

```
-------------+------------------------------------------------
greg_distn~t |
           0 |   91991.85   948.8083      90132.08    93851.62
           1 |   45348.52   847.8992      43686.54     47010.5
--------------------------------------------------------------
```

Mean Estimates for Ethnic Border Distance and Onset of Territorial Conflict

```
Mean estimation                     Number of obs    =   95886

            0: teronset = 0
            1: teronset = 1

                (Std. Err. adjusted for 15981 clusters in x_y)
--------------------------------------------------------------
             |               Robust
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
greg_distn~t |
           0 |   89839.85   918.2058      88040.07    91639.64
           1 |   127892.3   3613.335      120809.8    134974.9
--------------------------------------------------------------
```

Mean Estimates for Ethnic Border Distance and Occurrence of Governmental Conflict

```
Mean estimation                     Number of obs    =   95886

            0: govoccur = 0
```

```
          1: govoccur = 1

             (Std. Err. adjusted for 15981 clusters in x_y)
--------------------------------------------------------------
             |             Robust
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
greg_distn~t |
           0 |    92701.3   997.8793      90745.35    94657.26
           1 |    76414.3   1645.602      73188.74    79639.87
--------------------------------------------------------------
```

Mean Estimates for Ethnic Border Distance and Occurrence of Territorial Conflict

```
Mean estimation                    Number of obs    =   95886

          0: teroccur = 0
          1: teroccur = 1

             (Std. Err. adjusted for 15981 clusters in x_y)
--------------------------------------------------------------
             |             Robust
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
greg_distn~t |
           0 |    91662.6   944.0645      89812.13    93513.08
           1 |   75337.36   2136.724      71149.14    79525.58
--------------------------------------------------------------
```

Mean Estimates for Ethnic Border Density and Onset of Territorial Conflict

```
Mean estimation                         Number of obs    =   95886

            0: teronset = 0
            1: teronset = 1

                (Std. Err. adjusted for 15981 clusters in x_y)
--------------------------------------------------------------
             |             Robust
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
gre~ynocoast |
           0 |   1.210885   .0082166       1.19478    1.226991
           1 |    .922387   .0244368      .8744882    .9702859
--------------------------------------------------------------
```

Mean Estimates for Ethnic Border Density and Onset of Governmental Conflict

```
Mean estimation                         Number of obs    =   95886

            0: govonset = 0
            1: govonset = 1

                (Std. Err. adjusted for 15981 clusters in x_y)
--------------------------------------------------------------
```

```
             |               Robust
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
gre~ynocoast |
           0 |   1.196988   .0082232      1.180869    1.213106
           1 |   1.474502   .0186586      1.437929    1.511075
--------------------------------------------------------------
```

Mean Estimates for Ethnic Border Density and Occurrence of Territorial Conflict

```
Mean estimation                     Number of obs    =   95886

            0: teroccur = 0
            1: teroccur = 1

--------------------------------------------------------------
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
gre~ynocoast |
           0 |   1.199616   .0034901      1.192775    1.206456
           1 |   1.287272   .0116982      1.264344    1.310201
--------------------------------------------------------------
```

Mean Estimates for Ethnic Border Density and Occurrence of Governmental Conflict

```
Mean estimation                     Number of obs    =   95886
```

```
            0: govoccur = 0
            1: govoccur = 1

                (Std. Err. adjusted for 15981 clusters in x_y)
--------------------------------------------------------------
             |                Robust
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
gre~ynocoast |
           0 |   1.222376   .0087155      1.205293     1.23946
           1 |   1.099191   .0147425      1.070294    1.128088
--------------------------------------------------------------
```

Mean Estimates for Ethnic Group Configuration Percentage within the State and Onset of Governmental Conflict

```
Mean estimation                     Number of obs    =   95886

            0: govonset = 0
            1: govonset = 1

               (Std. Err. adjusted for 304 clusters in c_year)
--------------------------------------------------------------
             |                Robust
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
ratio_g1g2~t |
           0 |   .3201613    .015892      .2888887     .351434
           1 |   .2080859   .0663665      .0774883    .3386835
```

```
------------------------------------------------------------
```

Mean Estimates for Ethnic Group Configuration Percentage within the State and Onset of Territorial Conflict

```
Mean estimation                    Number of obs    =   95886

            0: teronset = 0
            1: teronset = 1

               (Std. Err. adjusted for 304 clusters in c_year)
--------------------------------------------------------------
             |               Robust
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
ratio_g1g2~t |
           0 |   .3179734   .0160494       .286391    .3495559
           1 |   .2400609   .0241932       .192453    .2876688
--------------------------------------------------------------
```

Mean Estimates for Ethnic Group Configuration Percentage within the State and Occurrence of Governmental Conflict

```
Mean estimation                    Number of obs    =   95886

            0: govoccur = 0
            1: govoccur = 1

               (Std. Err. adjusted for 304 clusters in c_year)
```

```
--------------------------------------------------------------
             |                Robust
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
ratio_g1g2~t |
           0 |   .3205108   .0162805      .2884738    .3525479
           1 |   .2914492   .0437223      .2054113     .377487
--------------------------------------------------------------
```

Mean Estimates for Ethnic Group Configuration Percentage within the State and Occurrence of Territorial Conflict

```
Mean estimation                    Number of obs    =   95886

            0: teroccur = 0
            1: teroccur = 1

               (Std. Err. adjusted for 304 clusters in c_year)
--------------------------------------------------------------
             |                Robust
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
ratio_g1g2~t |
           0 |   .3251812   .0165064      .2926995    .3576628
           1 |   .2036399   .0160875      .1719825    .2352972
--------------------------------------------------------------
```

Conclusion

This chapter has helped to weed through the sizeable number of variables into a set of likely candidates for further analysis in combined regression models. On the economic front, means analysis suggests that conflict is more likely as higher levels of rain experienced in an area. As a matter of modeling, this finding is probably partially derived from the relationship between rainfall levels and population, but the finding that increased rates of rainfall lead to conflict suggests that some more complicated causal factors may be involved. It also suggests that the common argument about resource scarcity is unlikely to hold at the local level. Similar to rainfall, growth in GDP appears to have a strong relationship with conflict, while other commonly associated variables, such as GDP per capita, did not. This finding may not deter further examination of GDP at the local level, but it reinforces the need for careful consideration of every variable at the local level, as even the most commonly accepted variables appear to be having interesting results. The combined finding that higher levels of rain and lower levels of economic growth are likely to lead to conflict suggests that those regions with the greatest inherent resources (from rain) but still not experiencing growth leads to a mismatch that can cause conflict. The rainfall finding is furthered when horizontal inequality is taken into consideration – it is the rainiest parts of countries that are apparently most prone to conflict. Again, this finding suggests that the "prize" theory of conflict location – that conflicts will center around territory of greatest value – may be justified by these statistics. The multiple variables tested for natural resource data also suggest that a middle approach focused on "prizes" is necessary for evaluating conflict risk: while distance to oil increases the risk of conflict, density of oil does not, suggesting the smaller deposits may make more easily captured targets for rebellion. On the non-fuel mineral front, it appears that only minerals that have extremely high value for weight – diamonds – are likely causes of conflict, while the presence of minerals in most other cases increases government presence and local development, thereby decreasing conflict risk. That conflict appears to be located in less unequal countries based on Gini factors suggests that conflict is made possible by a modicum of balance between different groups – in very unequal countries, it may be most difficult for resource-poor groups to mount an effective fight. The overall finding that the extremes of variables are not the most dangerous carries to terrain analysis as well, where hills and middle forest densities appeared to be more prone to produce conflict than the extreme values for mountains and dense forest. In the capital distance analysis, the consistent finding that distance from a capital increases the risks for territorial (but not government) conflict onset held while some of the consistent weaknesses in border data were also consistent. The population data also reinforces the difference between government and territorial conflict, with government conflict much more likely in populated areas while territorial conflict continues to hold to less populated places.

This overall approach – using a combination of local and state level variables – will be continued in the next chapter where the findings from this means analysis will be used to determine regression models that explain conflict. The general hypotheses that have been partially proven by this study of means differences – that local conflict tends towards more balanced situations and “prizes” – will be further examined in the regressions, hopefully leading to a combined regression result that is robust to different specifications.

Bibliography

50-Year Reconstruction. 2008. NOAA / National Weather Service 2008 [cited 17 November 2008]. Available from ftp://ftp.cpc.ncep.noaa.gov/precip/50yr/gauge/2.5deg/.

Alvarez, Maria D. 2003. "Forests in a Time of Violence: Conservation Implications of the Colombian War." *Journal of Sustainable Forestry* 16 (3 & 4):47-68.

Amey, Earle. 2000. "Gold." In *US Geological Survey Mineral Yearbook*: US Geological Survey.

Anselin, Luc, Jacqueline Cohen, David Cook, Wilpen Gorr, and George Tita. 2000. "Spatial Analyses of Crime." *Criminal Justice* 4:213-62.

Arnold, Guy. 1995. *Wars in the Third World Since 1945*. London: Continuum International Publishing Group.

Aspinall, Edward. 2007. "The Construction of Grievance: Natural Resources and Identity in a Separatist Conflict." *Journal of Conflict Resolution* 51 (6):950-72.

Ayalew, Lulseged, and Hiromitsu Yamagishi. 2005. "The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan." *Geomorphology* 65:15-31.

Baller, Robert D., Luc Anselin, Steven F. Messner, Glenn Deane, and Darnell F. Hawkins. 2001. "Structural Covariates of U.S. County Homicide Rates: Incorporating Spatial Effects." *Criminology* 39 (3):561-90.

Berridge, G. R. 1989. "Diplomacy and the Angola/Namibia Accords." *International Affairs* 65 (3):463-79.

Besley, Timothy, and Torsten Persson. 2009. "The Incidence of Civil War: Theory and Evidence." In *Economic Organisation and Public Policy Discussion Paper*. London: Suntory and Toyota International Centres for Economics and Related Disciplines, London School of Economics and Political Science.

Blattman, Christopher, and Edward Miguel. 2010. *Civil War: A Review of Fifty Years of Research*. Center for Global Development, March 2009 [cited 30 August 2010]. Available from www.cgdev.org/files/1421335_file_Blattman_Civil_War_FINAL.pdf.

Bloom, David E., and Jeffrey D. Sachs. 1998. "Geography, Demography and Economic Growth in Africa." In *Brookings Papers on Economic Activity*. Washington D.C.: Brookings Institution Press.

Boulding, Kenneth. 1962. *Conflict and Defense*. New York: Harper & Row.

Brandsegg, Kristian B. 2007. "Prospective oil and gas resources and its relation to potential insurgencies: New insurgency impending in South Eastern Sudan?" In *15th Annual National Political Science Conference*. Trondheim.

Bremmer, Ian. 2006. *The J Curve: A New Way to Understand Why Nations Rise and Fall*. New York: Simon and Schuster.

Bruk, Salomon I. 1964. *Atlas Naradov Mira*. Moscow: Academy of Science USSR.

Buhaug, Halvard. 2007. "Dude, Where's My Conflict? LSG, Relative Strength, and the Location of Civil War." In *International Studies Association Annual Convention*. Chicago.

Buhaug, Halvard, Lars-Erik Cederman, and Jan Ketil Rød. 2006. "Modeling Ethnic Conflict in Center-Periphery Dyads." In *Polarization and Conflict Workshop*. Nicosia, Cyprus.

———. 2009. "Disaggregating Ethno-Nationalist Civil Wars: A Dyadic Test of Exclusion Theory."

Buhaug, Halvard, and Scott Gates. 2002. "The Geography of Civil War." *Journal of Peace Research* 39 (4):417-33.

Buhaug, Halvard, Scott Gates, and Päivi Lujala. 2005. "Geography, Strategic Ambition, and the Duration of Civil Conflict." In *Mapping the Complexity of Civil Wars*. Zurich.

Buhaug, Halvard, and Kristian Gleditsch. 2005. "The Origin of Conflict Clusters: Contagion or Bad Neighborhoods?" In *Third European Consortium for Political Research General Conference*. Budapest, Hungary.

Buhaug, Halvard, and Päivi Lujala. 2005. "Accounting for Scale: Measuring Geography in Quantitative Studies of Civil War." *Political Geography* 24:399-418.

Buhaug, Halvard, and Jan Ketil Rød. 2006. "Local Determinants of African Civil Wars, 1970-2001." *Political Geography* 25:315-35.

Callaway, H.M., Gertrude N. Greenspoon, and Wilma F. Washington. 1961. "Copper." In *Minerals yearbook metals and minerals (except fuels)*: Bureau of Mines, US Geological Survey.

Carey, Sabine C. 2007. "Rebellions in Africa: Disaggregating the Effect of Political Regimes." *Journal of Peace Research* 44 (1):47-64.

Cederman, Lars-Erik. 2003. "Modeling the Size of Wars: From Billiard Balls to Sandpiles." *American Political Science Review* 97 (1):135-50.

Cederman, Lars-Erik, Halvard Buhaug, and Jan Ketil Rød. 2009. "Ethno-Nationalist Dyads and Civil War: A GIS-Based Analysis." *Journal of Conflict Resolution* 53 (4):496-525.

Cederman, Lars-Erik, Jan Ketil Rød, and Nils Weidmann. 2006. "Geo-Referencing of Ethnic Groups: Creating a New Dataset." In *GROW Workshop*. Oslo, Norway.

———. 2007. "Geo-Referencing Ethnic Groups: Creating a New Dataset." In *International Studies Association Annual Convention*. Chicago.

Chad: A Country Study. US Marine Corps [cited. Available from http://www.marines.mil/news/publications/Documents/Chad%20Study_3.pdf.

Clark, Jeffrey. 1993. "Debacle in Somalia." *Foreign Affairs* 72 (1):109-23.

Clausewitz, Carl von. 1976. *On War*. Translated by M. Howard and P. Paret. Edited by M. Howard and P. Paret. Princeton: Princeton University Press.

Clover, Jenny. 2003. "Food Security in Sub-Saharan Africa." *African Security Review* 12 (1):5-15.

Collier, Paul, and Jan Willem Gunning. 1999. "Explaining African Economic Performance." *Journal of Economic Literature* 37 (1):64.

Collier, Paul, and Anke Hoeffler. 1998. "On economic causes of civil war." *Oxford Economic Papers* 50:563-73.

———. 2002. "On the Incidence of Civil War in Africa." *Journal of Conflict Resolution* 46 (1):13-28.
———. 2004. "Greed and Grievance in Civil War." *Oxford Economic Papers* 56:563-95.
Collier, Paul, Anke Hoeffler, and Måns Söderbom. 2004. "On the Duration of Civil War." *Journal of Peace Research* 41 (3):253-73.
Collier, Paul, Patrick Honohan, and Karl Ove Moene. 2001. "Implications of Ethnic Diversity " *Economic Policy* 16 (32):129-66.
Cunningham, David, Kristian Gleditsch, and Idean Salehyan. 2005. "Dyadic Interactions and Civil War Duration." In *International Studies Association Annual Convention*. Honolulu, HI.
Denton, F.H. 1966. "Some Regularities in International Conflict, 1820-1949." *Background* 9 (4):283-96.
"Digital Chart of the World." 2007. Environmental Systems Research Institute.
Doyle, Michael W., and Nicholas Sambanis. 2000. "International Peacebuilding: A Theoretical and Quantitative Analysis." *American Political Science Review* 94 (4):778-801.
Duvall, Raymond, and Mary Wefling. 1973. "Social Mobilization, Political Institutionalization, and Conflict in Black Africa: A Simple Dynamic Model." *Journal of Conflict Resolution* 17 (4):673-702.
Elbadawi, Ibrahim, and Nicholas Sambanis. 2000. "Why Are There So Many Civil Wars in Africa? Understanding and Preventing Violent Conflict." *Journal of African Economies* 9 (3):244-69.
———. 2002. "How Much War Will We See? Explaining the Prevalence of Civil War." *Journal of Conflict Resolution* 46 (3):307-34.
Elkins, Zachary. 2000. "Gradations of Democracy? Empirical Tests of Alternative Conceptualizations." *American Journal of Political Science* 44 (2):293-300.
Ellingsen, Tanja, and Nils Petter Gleditsch. 1997. "Democracy and Armed Conflict in the Third World." In *Causes of Conflict in Third World Countries*, ed. K. Volden and D. Smith. Oslo: North-South Coalition and International Peace Research Institute, Oslo.
FAO. 2007. *Armed conflict, hunger threaten sustainable mountain development in Asia-Pacific, says FAO*, 13 May 2002 [cited 14 December 2007]. Available from http://www.fao.org/world/regional/rap/news_detail.asp?event_id=25641&year=2002.
———. 2005. *State of the World's Forests*. Rome: Food and Agriculture Organization of the United Nations.
Fearon, James D. 2003. "Ethnic and Cultural Diversity by Country." *Journal of Economic Growth* 8:195-222.
———. 2004. "Why Do Some Civil Wars Last So Much Longer Than Others?" *Journal of Conflict Resolution* 41 (3):275-301.
———. 2005. "Primary Commodity Exports and Civil War." *Journal of Conflict Resolution* 49 (4):483-507.
Fearon, James D., and David D. Laitin. 1999. Weak States, Rough Terrain, and Large Scale Ethnic Violence Since 1945. Paper read at Annual Meetings of the American Political Science Association, 2-5 September, at Atlanta, Georgia.

———. 2010. *Additional Tables for "Ethnicity, Insurgency, and Civil War"*, 6 February 2003a [cited 28 August 2010]. Available from www.stanford.edu/~jfearon/papers/addtabs.pdf.

———. 2003b. "Ethnicity, Insurgency, and Civil War." *American Political Science Review* 97 (1):75-90.

———. 2010. *Sons of the Soil, Migrants, and Civil War*. Stanford University, 2 August 2009 [cited 1 February 2010]. Available from http://www.stanford.edu/~jfearon/papers/SoS.doc.

Fotheringham, A. Stewart, Chris Brunsdon, and Martin Charlton. 2002. *Geographically Weighted Regression: the Analysis of Spatially Varying Relationships*. West Sussex, England: John Wiley & Sons Ltd.

"Fresh violence in Bolivia stokes civil war fears." 2008. *Agence France Presse*, 11 September.

Gates, Scott, Havard Hegre, Mark P. Jones, and Havard Strand. 2006. "Institutional Inconsistency and Political Instability: Polity Duration, 1800-2000." *American Journal of Political Science* 50 (4):893-908.

Gelman, Andrew, and Andrew Hill. 2006. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Edited by R. M. Alvarez, N. L. Beck and L. L. Wu. Cambridge: Cambridge University Press.

Gerges, Fawaz A. 2000. "The End of the Islamist Insurgency in Egypt?: Costs and Prospects." *Middle East Journal* 54 (4):592-612.

Gilmore, Elisabeth, Nils Petter Gleditsch, Päivi Lujala, and Jan Ketil Rod. 2005a. "Conflict Diamonds: A New Dataset." *Conflict Management and Peace Science* 22 (3):257-72.

Gilmore, Elisabeth, Päivi Lujala, Nils Petter Gleditsch, and Jan Ketil Rød. 2009. *The Diamond Dataset Codebook* 2005b [cited 10 February 2009]. Available from http://www.prio.no/sptrans/1122200499/DIADATA%20Codebook.pdf.

Giordano, Benito. 2000. "Italian regionalism or 'Padanian' nationalism — the political project of the Lega Nord in Italian politics " *Political Geography* 19 (4):445-71.

Gleditsch, Nils Petter, Havard Strand, Mikael Eriksson, Margareta Sollenberg, and Peter Wallensteen. 2001. Armed Conflict 1946–99: A New Dataset. Paper read at Identifying Wars: Systematic Conflict Research and Its Utility in Conflict Resolution and Prevention, 8-9 June, at Uppsala, Sweden.

Gleditsch, Nils Petter, Havard Strand, and Lars Wilhelmsen. 2002a. "Armed Conflict Dataset Codebook." Oslo: International Peace Research Institute.

Gleditsch, Nils Petter, Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg, and Havard Strand. 2002b. "Armed Conflict 1946-2001: A New Dataset." *Journal of Peace Research* 39 (5):615-37.

Global Forest Cover. 2010. United Nations Environment Program 2000 [cited 1 February 2010]. Available from http://geodata.grid.unep.ch.

"Global Resource Information Database." 2007. United Nations Environment Program.

Global Terrorism Database. 2010. *Global Terrorism Database*. National Consortium for the Study of Terrorism and Responses to Terrorism, University of Maryland 2010a [cited 4 October 2010]. Available from http://www.start.umd.edu/gtd/search/Results.aspx?page=2&casualties_type=b&casualties_max=&start_yearonly=1970&end_yearonly=1999&dtp2=all&region=11&c

ount=100&expanded=no&charttype=line&chart=overtime&ob=TotalNumberOfFatalities&od=desc#results-table.

———. 2010. *Incident Summary for GTDID: 199810250005*. National Consortium for the Study of Terrorism and Responses to Terrorism, University of Maryland 2010b [cited 4 October 2010]. Available from http://www.start.umd.edu/gtd/search/IncidentSummary.aspx?gtdid=199810250005.

Global Witness. 2002. "The Logs of War." In *Programme for International Co-operation and Conflict Resolution*. Oslo: Fafo Institute for Applied Social Science.

———. 2005. "Under-Mining Peace: Tin - the Explosive Trade in Cassiterite in Eastern DRC." London: Global Witness.

———. 2006. "Digging in Corruption." London: Global Witness.

Goldstone, Jack A., Robert H. Bates, David L. Epstein, Ted Robert Gurr, Michael B. Lustik, Monty G. Marshall, Jay Ulfelder, and Mark Woodward. 2010. "A Global Model for Forecasting Political Instability." *American Journal of Political Science* 54 (1):190-208.

Grimes, Barbara. 2000. *Ethnologue: Languages of the World*. 14th ed. Dallas, T.X.: SIL International.

Gritzner, Mandy Lineback, W. Andrew Marcus, Richard Aspinall, and Stephan G. Custer. 2001. "Assessing landslide potential using GIS, soil wetness modeling and topographic attributes, Payette River, Idaho." *Geomorphology* 37:149-65.

Gundel, Joakim. 2002. "The Migration–Development Nexus: Somalia Case Study." *International Migration* 40 (5):255-81.

Gurr, Ted. 1968. "A Causal Model of Civil Strife: A Comparative Analysis Using New Indices." *American Political Science Review* 62 (4):1104-24.

———. 1970. *Why Men Rebel*. Princeton, N.J.: Princeton University Press.

———. 1974a. "Persistence and Change in Political Systems, 1800-1971." *American Political Science Review* 68 (December):1482-504.

Gurr, Ted Robert. 1974b. "Persistence and Change in Political Systems, 1800-1971." *American Political Science Review* 68 (3):December.

Harbom, Lotta. 2010. *UCDP/PRIO Armed Conflict Dataset Codebook, Version 4-2009*. Uppsala Conflict Data Program (UCDP), 17 June 2009 2009 [cited 4 October 2010]. Available from http://www.prio.no/sptrans/-1423485763/Codebook_UCDP_PRIO%20Armed%20Conflict%20Dataset%20v4_2009.pdf.

Haring-Smith, Whitney. 2005. "PLSC 195b: Civil War Final Paper: Economic and Military Interpretations of GDP per capita in Models of Civil War Risk." Yale University.

———. 2007. "Coding of Ethnic Wars at the Local and National Level: An Examination of the Chittagong Hills Case and its Implications for our Empirical Understanding of Ethnic War." New Haven: Yale University.

———. 2008. 'All Conflict is Local': An Empirical Analysis of Local Factors in Violent Civil Conflict. Paper read at Probationary Research Student Seminar, 8 February, at Oxford, UK.

Hauge, Wenche, and Tanja Ellingsen. 2001. "Causal Pathways to Conflict." In *Environmental Conflict*, ed. P. Diehl and N. P. Gleditsch. Boulder, CO: Westview.

Hegre, Havard. 2002. Some Social Requisites of Democratic Peace: Democracy, Development, and Armed Conflict. Paper read at American Political Science Association Annual Meeting, 29 August - 1 September, at Boston, MA.
Hegre, Havard, Tanja Ellingsen, Scott Gates, and Nils Petter Gleditsch. 2001a. "Toward a Democratic Civil Peace? Democracy, Political Change, and Civil War, 1816–1992." *American Political Science Review* 95 (1):33-48.
———. 2001b. "Toward a Democratic Civil Peace? Democracy, Political Change, and Civil War, 1916-1992." *American Political Science Review* 95 (March):33-49.
Hegre, Havard, Gudrun Østby, and Clionadh Raleigh. 2007. Economic Deprivation and Civil War Events: A Disaggregated Study of Liberia. Paper read at ISV, 25 March, at Tirsdagsseminaret.
Hegre, Havard, and Clionadh Raleigh. 2005. Population Size, Concentration, and Civil War, a Geographically Disaggregated Analysis. Paper read at Mapping the Complexity of Civil Wars, 6 September, at Zurich.
Hegre, Havard, and Nicholas Sambanis. 2006. "Sensitivity Analysis of Empirical Results on Civil War Onset." *Journal of Conflict Resolution* 50 (4):508-35.
Hendrix, Cullen S. 2006. "“Head for the Hills!” or This Hard Land? The Direct and Indirect Effects of Rough Terrain on Civil Conflict Onset." In *International Studies Association Annual Convention*. San Diego, CA.
Herbst, Jeffrey. 2000. *States and power in Africa*. Princeton, N.J.: Princeton University Press.
Hoeffler, Anke. 2009. "On the Causes of Civil War (draft)." In *Oxford Handbook of the Economics of Peace and Conflict*, ed. M. Garfinkel and S. Skaperdas. Oxford, UK: Oxford University Press.
Horowitz, Donald L. 2000. *Ethnic Groups in Conflict*. Berkeley: University of California Press.
Howard, Michael. 1989. *War and the Liberal Conscience (The Trevelyan lectures)*. Oxford: Oxford University Press.
Humphreys, Macartan. 2005. "Natural Resources, Conflict, and Conflict Resolution: Uncovering the Mechanisms." *Journal of Conflict Resolution* 49 (4):508-38.
Humphreys, Macartan, and Habaye Ag Mohamed, eds. 2005. *Senegal and Mali*. Edited by P. Collier and N. Sambanis. Vol. 1. Washington, D.C.: The World Bank.
Humphreys, Macartan, and Jeremy M. Weinstein. 2006. "Handling and Manhandling Civilians in Civil War." *American Political Science Review* 100 (3):429-47.
Huntington, Samuel. 1991. *The Third Wave: Democratization in the Late Twentieth Century*. Norman, OK and London: University of Oklahoma Press.
———. 1993. "The Clash of Civilizations?" *Foreign Affairs* 72 (3):44-57.
———. 1996. *The Clash of Civilizations and the Remaking of World Order*. New York: Simon and Shuster.
Jackson, Stephen. 2003. "Fortunes of war: the coltan trade in the Kivus." In *Power, livelihoods, and conflict: case studies in political economy analysis for humanitarian action*, ed. S. Collinson. London: Overseas Development Institute.
Jaggers, Keith, and Ted Robert Gurr. 1995. "Tracking Democracy's Third Wave with the Polity III Data." *Journal of Peace Research* 32:469-82.
Jensen, Peter Sandholdt and Kristian Gleditsch. 2008. "Rain, Growth, and Civil War: The Importance of Location." Human Security Gateway.

Kahl, Colin. 2006. "Plight or Plunder? Natural Resources and Civil War." Pittsburgh, PA: Ridgway Working Group on the Political Economy of International Security.
Kaldor, Mary, Terry Karl, and Yahia Said, eds. 2007. *Oil Wars*. London: Pluto Press.
Keita, Kalifa. 1998. "Conflict and conflict resolution in the Sahel: The Tuareg insurgency in Mali." *Small Wars & Insurgencies* 9 (3):102-28.
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):145-66.
———. 2010. *The Battle Deaths Dataset, 1946−2005, Version 2.0: Documentation of Coding Decisions*. Centre for the Study of Civil War (CSCW), International Peace Research Institute (PRIO), September 2006 [cited 8 January 2010]. Available from http://www.prio.no/misc/Download.aspx?file=%2fprojects%2fdataset-website-workspace%2fBattle%2520Deaths%2520Data%252020%2fDocumentation_PRIO-UCDP.pdf.
Landau, Daniel. 1986. "Government and Economic Growth in the Less Developed Countries: An Empirical Study for 1960-1980." *Economic Development and Cultural Change* 35 (1):35-75.
Lawrence, T.E. 2007. *Science of Guerilla Warfare* 1929 [cited 15 December 2007]. Available from http://www.bellum.nu/literature/lawrence001.html.
Le Billon, Philippe. 2001a. "Angola's Political Economy of War: The Role of Oil and Diamonds, 1975-2000." *African Affairs* 100:55-80.
———. 2001b. "The Political Ecology of War: Natural Resources and Armed Conflicts." *Political Geography* 20:561-84.
———. 2005. "Resources and Armed Conflicts." *The Adelphi Papers* 45 (373).
Lewis, I. M. 1989. "The Ogaden and the Fragility of Somali Segmentary Nationalism." *African Affairs* 88 (353):573-9.
Lewis, Paul M. 2009. *Ethnologue: Languages of the World, Sixteenth edition*. Dallas: SIL International.
Liang, Jun. 2001. Simulating Crimes and Crime Patterns Using Cellular Automata and GIS, Geography, University of Cincinnati, Cincinnati.
Lujala, Päivi. 2005. "The Spoils of Nature: Armed Civil Conflict and Rebel Access to Natural Resources."
———. 2010. "The spoils of nature: Armed civil conflict and rebel access to natural resources." *Journal of Peace Research* 47 (1):15-28.
Lujala, Päivi, Nils Petter Gleditsch, and Elisabeth Gilmore. 2005. "A Diamond Curse?" *Journal of Conflict Resolution* 49 (4):538-62.
Lujala, Päivi, Jan Ketil Rød, and Nadja Thieme. 2007. "Fighting Over Oil: Introducing a New Dataset." *Conflict Management and Peace Science* 24 (3):239-56.
Machiavelli, Niccolo. 1520. "Art of War."
Malaquias, Assis. 2001. "Diamonds are a guerrilla's best friend: the impact of illicit wealth on insurgency strategy." *Third World Quarterly* 22 (3):311-25.
Marshall, Monty G., and Keith Jaggers. 2007a. "Polity IV Annual Time-Series, 1800-2007." Center for Systemic Peace.
———. 2009. *Polity IV Project Data Users' Manual*. Center for Systemic Peace 2007b [cited 23 February 2009]. Available from http://www.systemicpeace.org/inscr/p4manualv2007.pdf.

Menkhaus, Ken. 2007. "Governance without Government in Somalia: Spoilers, State Building, and the Politics of Coping." *International Security* 31 (3):74-106.
Miguel, Edward and Shanker Satyanath and Ernest Sergenti. 2004. "Economic Shocks and Civil Conflict: An Instrumental Variables Approach." *Journal of Political Economy* 112 (4):725-53.
Mineral Resources Data System. 2009. United States Geological Survey 2005 [cited 9 February 2009]. Available from http://tin.er.usgs.gov/mrds/.
Montague, Dena. 2002. "Stolen Goods: Coltan and Conflict in the Democratic Republic of Congo." *SAIS Review* 22 (1):103-18.
Muller, Edward N., and Erich Weede. 1990. "Cross-National Variations in Political Violence: A Rational Action Approach." *Journal of Conflict Resolution* 34 (6):624-51.
Munck, Gerardo L., and Jay Verkuilen. 2002. "Conceptualizing and Measuring Democracy: Evaluating Alternative Indices." *Comparative Political Studies* 35 (February):5-34.
Murshed, S. Mansoob, and Scott Gates. 2003. Spatial-Horizontal Inequality and the Maoist Insurgency in Nepal. Paper read at UNU/WIDER Project Conference on Spatial Inequality in Asia, 28-29 March, at United Nations University, Tokyo.
Ohlmacher, Gregory C., and John C. Davis. 2003. "Using multiple logistic regression and GIS technology to predict landslide hazard in northeast Kansas, USA." *Engineering Geology* 69:331-43.
Olsson, Ola. 2005. "Conflict diamonds." *Journal of Development Economics* 82:267-86.
Olsson, Ola, and Heather Congdon Fors. 2004. "Congo: The Prize of Predation." *Journal of Peace Research* 41 (3):321-36.
Osaghae, Eghosa E. 1995. "The Ogoni Uprising: Oil Politics, Minority Agitation, and the Future of the Nigerian State." *African Affairs* (94):325 - 44.
Østby, Gudrun. 2007. "Horizontal Inequalities, Political Environment, and Civil Conflict, Evidence from 55 Developing Countries 1986–2003." In *Post-Conflict Transitions Working Paper, No. 7*. Washington, DC: World Bank.
Parker, Karen, Anne Heindel, and Adam Branch. 2010. *Armed Conflict in the World Today: A Country by Country Review*. Humanitarian Law Project 2000 [cited 9 October 2010]. Available from hlp.home.igc.org/docs/word_docs/Armed_Conflict_2000.doc.
Peceny, Mark, Caroline Beer, and Shannon Sanchez-Terry. 2002. "Dictatorial Peace?" *American Political Science Review* 96 (2):15-26.
Prunier, Gérard. 2004. "Rebel movements and proxy warfare: Uganda, Sudan, and the Congo (1986-99)." *African Affairs* 103 (412):359-83.
Raleigh, Clionadh, and Havard Hegre. 2005. Introducing ACLED: An Armed Conflict Location and Event Dataset. Paper read at Disaggregating the Study of Civil War and Transnational Violence, 7-8 March, at San Diego, CA.
———. 2009. "Population size, concentration, and civil war. A geographically disaggregated analysis." *Political Geography* 28:224-38.
Rapoport, Anatol. 1957. "Lewis F. Richardson's Mathematical Theory of War." *Conflict Resolution* 1 (3):249-99.
Regan, Patrick. 1996. "Conditions for Successful Third Party Interventions." *Journal of Conflict Resolution* 40 (1):336-59.

"Report of the Panel of Experts on the Illegal Exploitation of Natural Resources and Other Forms of Wealth of the Democratic Republic of the Congo." 2001. S/2001/357: United Nations.

Reuveny, Rafael, and John W. Maxwell. 2001. "Conflict and Renewable Resources." *Journal of Conflict Resolution* 45 (6):719-42.

Reynal-Querol, Marta. 2002. "Political Systems, Stability, and Civil War." *Defense and Peace Economics* 13 (6):465-83.

Richardson, Lewis F. 1948. "Variation of the Frequency of Fatal Quarrels with Magnitude." *American Statistical Association* 43:523-46.

———. 1960. *Statistics of Deadly Quarrels*. Chicago: Quadrangle Books.

Rød, Jan Ketil, and Halvard Buhaug. 2005. "Using disaggregated grid cells for a study on the onset of African Civil Wars."

———. 2007. Civil Wars: Prospects and Problems with the Use of Local Indicators. Paper read at Scandinavian Research Conference on Geographical Information Sciences, 5-7 September at Ås, Norway.

Roeder, Philip G. 2008. *Ethnolinguistic Fractionalization (ELF) Indices, 1961 and 1985* 2001 [cited 16 November 2008]. Available from http//:weber.ucsd.edu\~proeder\elf.htm.

Ross, Michael. 2002. "Oil, Drugs, and Diamonds: How Do Natural Resources Vary in their Impact on Civil War?" In *International Peace Academy Project Economic Agendas in Civil Wars*.

———. 2004a. "How Do Natural Resources Influence Civil War? Evidence from Thirteen Cases." *International Organization* 58:35-67.

———. 2004b. "What Do We Know about Natural Resources and Civil War." *Journal of Peace Research* 41 (5):337-56.

———. 2006. "A Closer Look at Oil, Diamonds, and Civil War." *Annual Review of Political Science* 9:265-300.

Rustad, Siri Aas. 2005a. Forest Resources and Conflict: How Forest Resources Influence Internal Armed Conflicts, Department of Sociology and Political Science, Norwegian University of Science and Technology, Trondheim, Norway.

———. 2005b. Forest Resources and Conflict: How Forest Resources Influence Internal Armed Conflicts, Norwegian University of Science and Technology (NTNU), Oslo.

Rustad, Siri Aas, Jan Ketil Rod, and Wenche Larsen. 2006. "Foliage and Fighting: A Skeptical View." Oslo, Norway: International Peace Research Institute.

Rustad, Siri Camilla Aas, Jan Ketil Rød, Wenche Larsen, and Nils Petter Gleditsch. 2008. "Foliage and fighting: Forest resources and the onset, duration, and location of civil war." *Political Geography* 27 (7):761-82.

Ryan, J.P., and Kathleen M. McBreen. 1961. "Gold." In *Minerals yearbook metals and minerals (except fuels)*: Bureau of Mines, US Geological Survey.

Sachs, Jeffrey. 2005. The End of Poverty: Economic Possibilities for Our Time, 30 March, at Carnegie Council.

Salehyan, Idean. 2010. "Transnational Insurgencies and the Escalation of Regional Conflict: Lessons for Iraq and Afghanistan." Carlisle, PA: Strategic Studies Institute, US Army War College.

Sambanis, Nicholas. 2001. "Do Ethnic and Nonethnic Civil Wars Have the Same Causes? A Theoretical and Empirical Inquiry (Part 1)." *Journal of Conflict Resolution* 45 (3):259-82.
———. 2003. "Using Case Studies to Expand the Theory of Civil War." *CPR Working Papers* (5).
———. 2010. *Civil War Coding Notes*. Yale University, 30 November 2004 [cited 8 October 2010]. Available from http://pantheon.yale.edu/%7Ens237/index/research/CivilWarCodingNotes.pdf.
Sanhueza, Ricardo. 1999. "The Hazard Rate of Political Regime." *Public Choice* 98 (June):337-67.
Silverman, B.W. 1986. *Density Estimation for Statistics and Data Analysis*. New York: Chapman and Hall.
Singer, J.D. 1987. "Reconstructing the Correlates of War Dataset on Material Capabilities of States, 1816-1985." *International Interactions* 14:115-32.
Singer, J.D., and Paul F. Diehl, eds. 1990. *Measuring the Correlates of War*. Ann Arbor, MI: University of Michigan Press.
Singer, J.D., and Melvin Small. 1982. *Resort to Arms: International and Civil War, 1816-1980*. Beverly Hills, CA: Sage.
———. 1994. "Correlates of War Project: International and Civil War Data, 1816–1992." ICPSR.
Sinha, Kounteya. 2009. "As panic rises, NICD lab deluged with samples." *Times of India*, 9 August.
Smillie, Ian, Lansana Gberie, and Ralph Hazleton. 2000. *The Heart of the Matter: Sierra Leone, Diamonds, and Human Security*. Darby, PA: DIANE Publishing.
Snyder, Richard, and Ravi Bhavnani. 2005. "Diamonds, Blood, and Taxes: A Revenue-Centered Framework for Explaining Political Order." *Journal of Conflict Resolution* 49 (4):563-97.
Stewart, Frances. 2002. "Horizontal Inequalities: A Neglected Dimension of Development." In *CRISE Working Paper No. 1*: Center for Research on Inequality, Human Security, and Ethnicity, Oxford University.
Sutherland, Ian. 1962. "Statistics of Human Conflict." *Journal of the Royal Statistical Society* 125 (3):473-83.
Theisen, Ole Magnus. 2006. Renewable Resource Scarcity and Internal Armed Conflict: A Quantitative Reassessment, Center for the Study of Civil War, PRIO, Oslo.
Toft, Monica Duffy. 2003. *The Geography of Ethnic Violence: Identity, Interests, and the Indivisibility of Territory*. Princeton, NJ: Princeton University Press.
———. 2007. "Population Shifts and Civil War: A Test of Power Transition Theory." *International Interactions* 33 (3):243-69.
Treier, Shawn, and Simon Jackman. 2008. "Democracy as a Latent Variable." *American Journal of Political Science* 51 (1):201-17.
Trunfino, Giuseppe. 2004. Predicting Wildfire Spreading Through a Hexagonal Cellular Automata Model. Paper read at 6th International Conference on Cellular Automata for Research and Industry, October, at Amsterdam, The Netherlands.
Tzu, Sun. "Art of War."

U.S. Geological Survey. 2010. *GTOPO30*. U.S. Department of the Interior, 20 October 2009 [cited 1 February 2010]. Available from http://eros.usgs.gov/#/Find_Data/Products_and_Data_Available/gtopo30_info.

UNU-WIDER World Income Inequality Database, Version 2.0c. 2008. United Nations University 2008 [cited 17 November 2008]. Available from http://www.wider.unu.edu/research/Database/en_GB/database/.

Urdal, Henrik. 2008. "Population, Resources, and Political Violence: A Subnational Study of India, 1956-2002." *Journal of Conflict Resolution* 52 (4):590-617.

Vesperini, Helen. 2009. *Congo's coltan rush*. BBC News, 1 August 2001 [cited 9 February 2009]. Available from http://news.bbc.co.uk/2/hi/africa/1468772.stm.

Vreeland, James. 2008. "The Effect of Political Regime on Civil War: Unpacking Anocracy." *Journal of Conflict Resolution* 52 (3).

Wallensteen, Peter, and Margareta Sollenberg. 2001. "Armed Conflict, 1989-2000." *Journal of Peace Research* 38 (5):629-44.

Weidmann, Nils. 2008a. Violence and the Changing Ethnic Map: the Interdependence of Territory and Conflict in Bosnia. Paper read at 49th Annual Convention of the International Studies Association, 26-29 March, at San Francisco, CA.

Weidmann, Nils, Havard Hegre, and Clionadh Raleigh. 2006. "Modeling Spatial and Temporal Patterns of Civil War." In *American Political Science Association Annual Meeting*. Philadelphia, PA.

Weidmann, Nils, Jan Ketil Rød and Lars-Erik Cederman. 2008b. "Representing Ethnic Groups in Space: A New Dataset." Zurich, Switzerland: ETH Zurich.

Weidmann, Nils, and Michael Ward. 2010. "Predicting Conflict in Space and Time." *Journal of Conflict Resolution* 54 (6):883-901.

Weinstein, Jeremy M. 2007. *Inside Rebellion: the Politics of Insurgent Violence*. New York, NY: Cambridge University Press.

Witness, Global. 2009. *The Truth About Diamonds: Conflict and Development* 2006 [cited October 3 2009]. Available from http://www.globalwitness.org/data/files/media_library/7/en/the_truth_about_diamonds.pdf.

Wrigley, Neil, Tim Holt, David Steel, and Mark Tranmer. 1997. "Analysing, modelling, and resolving the ecological fallacy." In *Spatial Analysis: Modelling in a GIS Environment*, ed. P. Longley and M. Batty. New York: John Wiley & Sons, Inc.

Zedong, Mao. 1937. "Guerilla Warfare."