Geography, Outcome, and Casualties: A Unified Model of Insurgency

Sebastian Schutte / International Conflict Research / ETH Zurich*†

Nations are born, grow, fight, conquer, are conquered, become empires, and rise and fall on the great stage of physical geography and human passions and knowledge, and not on the homogeneous white planes on which we draw our diagrams.
Kenneth Boulding

Abstract

This study introduces a theoretical model of how insurgency develops as a function of reactive mobilization. The theory extends a classic distance-decay model (Boulding 1962) by incorporating Kalyvas' (2006) typology of violence. It implies that geographic conditions crucially determine the accuracy of applied violence and thereby its public perception, which in turn determines the actors' ability to mobilize. As a first test of these effects, I propose a new geographic indicator that expresses the spatial accessibility of a country's population for both central governments and peripheral insurgent movements. Two empirical implications of the theory are tested with a large-N dataset on outcomes and casualties in insurgencies. The new indicator is significantly associated with both military outcomes and the number of casualties in insurgencies since 1970 and strengthens statistical predictions.

*I would like to thank Lars-Erik Cederman and Camber Warren for helpful comments on multiple versions of this paper. Elin Arbin gave valuable linguistic advice. Nils Bormann, Julian Wucherpfennig, Christa Deiwiks, and Corinne Bara provided valuable feedback. Han Droussen and Nils Weidmann were very helpful as discussants of earlier versions of this paper presented at the 2011 ECPR conference in Reykjavik and the 2011 workshop on "Inequality, Grievances and Civil War" in Zurich, Switzerland. Halvard Buhaug, Tobias Böhmelt, and other participants of the 2011 Jan Tinbergen Conference in Amsterdam also contributed helpful comments. Finally, Livia Schubiger provided very substantial feedback.
†schuttes@ethz.ch

1 Introduction

The bloody stalemate in Afghanistan, a highly volatile security situation in Iraq, and the different fates of the uprisings of the Arab Spring remind us that civil war outcomes and severities are difficult to predict. Most civil conflicts in the post-WWII era share a common type of warfare: insurgency (Fearon and Laitin, 2003, Kalyvas, 2005, Fall, 1965, Kalyvas and Balcells, 2010).[1] The majority of quantitative research on civil war is focused on war onset and initial motivations of the rebels. However, civil wars can display puzzling attributes once they are underway, including specific spatial distributions of violence (Buhaug and Gates, 2002, Allen, 2003, Raleigh and Hegre, 2009, Buhaug and Gleditsch, 2008, Kalyvas and Kocher, 2009, O'Loughlin and Witmer, 2010) and counterintuitive mobilization dynamics: Insurgent movements can start with a very small number of initial combatants and grow in size despite heavy losses and military inferiority to the incumbent (Fall, 1965, Macaulay, 1978, Herbst, 2004). It is therefore obvious that a more complete understanding of civil wars must also examine their internal processes and drivers, as has been done in a rapidly growing number of recent publications (Lyall and Wilson, 2009, Buhaug, 2010, Weidmann, 2011, Berman et al., 2011). This paper attempts to contribute to this line of research by providing a micro-level explanation for how geographic conditions affect outcomes and casualties in insurgencies. Drawing on Boulding's (1962) Loss of Strength Gradient and Kalyvas' (2006) typology of violence, a distance-decay mechanism for the accuracy of violence in insurgencies is proposed. Violence projected over short distances can be applied more selectively and conditionally than violence projected over long distances. Consequently, long-distance violence also causes a larger backlash in public perception and triggers reactive mobilization against the attacker. This theory implies that the outcomes of insurgencies crucially depend on the military actors' proximity to the bulk of the population. Moreover, scenarios in which the military actors are remote from the bulk of the population are likely to bring about the most destructive conflicts due to both sides' inaccurate violence, which in turn causes reactive mobilization.

The relevance of this project is twofold. Most importantly, a theoretical dispute on the effects of force is made explicit and resolved theoretically. Second, an empirical contribution is made in terms of a new

[1] To illustrate the importance of insurgency, consider the following numbers: Fearon and Laitin (2003) introduced a dataset with 127 cases of civil war since 1945 that led to at least 1,000 battle deaths. Drawing on a subset of this data, Connable and Libicki (2010) identified 89 clear cases of insurgency. However, Kalyvas and Balcells (2010) report a declining occurrence of insurgency in the post-Cold War era in comparison to symmetric unconventional confrontations in civil wars.

geographic indicator that allows us to test two observable implications of the theory regarding outcomes and casualties in insurgencies. While a complete test of the presented theory lies beyond the capabilities of a single paper, the empirical results on the macro-level are encouraging and can serve as a starting point for further research.

The paper proceeds as follows: A review of the existing literature in the next section paves the way for a subsequent theoretical discussion. After that, a theoretical synthesis between two influential, geographical models of armed conflict is derived. Two main hypotheses regarding outcomes and casualties based on the new model are tested in a large-N conflict-level analysis drawing on a newly defined geographical indicator. The results are in line with the theoretical expectations and strengthen statistical predictions. Finally, the implications of this study and possibilities for future research are discussed.

2 Existing Literature

Insurgency and irregular warfare received major attention throughout the twentieth century, starting with the works of Lawrence ([1927]1998) on the British-Arab guerrilla campaigns during Word War I. After World War II and the successful partisan activities against the Axis powers in Europe and South-East Asia, the topic received major attention in the works of communist theorists, such as Mao ([1938]1967) and Guevara (1961). The French defeat in Vietnam led to growing American involvement and counterinsurgency doctrine moved into the focus of Western military advisors and scholars (Galula, 1964, McColl, 1969). While the International Relations literature was mainly concerned with interstate war throughout much of the 1970s and 80s, the notion of insurgency has been increasingly emphasized in the last two decades.

There are two recurrent themes in all of these schools. First and foremost, every serious student and practitioner of insurgency has stressed its population-centric nature. In his famous metaphor, Mao ([1937]1961, 93) compared revolutionaries to the fishes swimming among the peasants as if they were the sea. He also stressed that a large population was necessary for guerrilla war to succeed (Mao [1938]1967, 11; 66). Guevara (1961, 2) echos this basic doctrine: “The guerrilla is supported by the peasant and worker masses of the region and of the whole territory in which it acts. Without these prerequisites, guerrilla warfare is not possible”. Consequently, the counterinsurgency school has been equally focused on the problem of public support.

Military advisor John Paul Vann was one of the first outspoken critics of military tactics that alienated the civilian population in Vietnam (Sheehan, 1988, 106). Awareness of the problem of civilian loyalties reached up to the highest levels of command. The expression "winning hearts and minds [of the Vietnamese]" was made famous by President Johnson. With the ongoing wars in Iraq and Afghanistan, this approach has reemerged and the emphasis on population-centric warfare is an essential part of today's doctrine (DOD, 2007, 1-27).

The question of why civilians participate in ongoing insurgencies has also been approached in conflict research. Some scholars portray the problem in terms of greed: Rebels engage in rebellion if the prospects of economic revenue exceed the individual risks. Following this logic, Grossman (1999, 269) states that in many insurgencies the rebels are indistinguishable from bandits or pirates, and Collier (2000) developed a model in which "rebellion is a special case of crime, with all the differences from conventional crime arising as a consequence of the particular loot sought by rebels" (page 841). Clearly, the solution to stopping an insurgency within this model is simple: Increase the risk for the rebels through the increased application of military force. Certain leaders have accepted this mechanism. As Kissinger put it: "I refuse to believe that a third-class power such as Vietnam does not have a breaking point" (Greiner, 2009, 22). Greiner coined the term "tonnage ideology" referring to the assumption that a sufficient amount of attrition will translate into enemy surrender through a breakdown of either will or risk-reward calculations (Greiner, 2009, 23).

The ratio of risks and rewards has been proposed as an explanation for why former civilians participate in insurgencies. According to the classic collective action problem (Olsen, 1965, Tullock, 1971, Lichback, 1995), increased risks should provide incentives for free-riding within the rebellion and therefore a weakening of its military capabilities. Translated into an observable implication, this approach suggests a positive association of the number of killed insurgents with the probability of incumbent victory. The strong emphasis on the "body count" as a measurement of military success in the Vietnam War is a prime example of this logic in action (Greiner, 2009, 105).

A series of scholarly and military contributions strongly disagrees with this theory and its implications. The use of indiscriminate violence in particular is assumed to have massively negative effects for the perpetrator (Mason and Krane, 1989, Kalyvas and Kocher, 2007b, Wood, 2003, Kalyvas, 2006, Condra and Shapiro, 2012, Braithwaite and Johnson, 2012, Linke et al., 2012, 150). The killing of innocent civilians and the destruction of property can alienate the population from the attacker. Even if violence is targeted against combatants, killing them might help the opponent to recruit additional combatants from their kin. General Stanley McChrystal refers to this effect as “insurgent math” (Hastings, 2010). Ellsberg (1970, 6) introduces an interesting metaphor. The rebels sometimes provoke military action by the state which in turn alienates the civilian population and aids the rebels. According to Ellsberg, this mechanism resembles the use of the opponent’s weight in judo: Instead of creating a comparative advantage for the attacker, the more weight they put behind their attack, the harder they fall themselves. Kilcullen (2009) coins the term “Accidental Guerrilla”, referring to elements of the local population that are drawn into the fight instead of being a priori adversaries of the incumbent. Clearly, this accidental process is closely linked to incumbent behavior in the field: As civilian casualties mount and destruction of property continues, more locals might be willing to retaliate independent of antecedent strategic loyalties. The US Army Counterinsurgency Handbook also stresses the importance of avoiding unnecessary destruction and violence and explicitly rejects the body count indicator (DOD, 2007, 5-27). While all of these accounts approach the problem from slightly different angles, the basic mechanism hereafter referred to as “reactive mobilization” is obvious: Instead of weakening the military opponent, violence can have the opposite effect. More civilians are repelled by the attacker and collaborate with the opponent.

Kalyvas (2006) provides a theoretically sound explanation of how violence against civilians is applied in civil wars. A central piece in this theory is the distinction between selective and indiscriminate violence. The former is applied more accurately against combatants, officers, and informants of the adversary, and the latter is applied less accurately against civilians. Drawing on a stylized spatial model that consists of five zones, Kalyvas (2006) concludes that selective violence is used mostly in zones of predominant but incomplete control, while indiscriminate violence is used if levels of control are low. With regard to how violence feeds back into the dynamics of control, the theory makes the case for both coercive and alienating effects, but ultimately assumes both violence and civilian collaboration to be largely endogenous to military control (Kalyvas, 2006, 12;118-132). In a recent study explicitly devoted to explaining the outcome of insurgencies, Lyall and Wilson (2009) argue that modern combat tactics have undermined the incumbent’s ability to win over the civilian population: Increasingly mechanized armies and long-range supply systems make contact with the civilian population less frequent and less important, thereby preventing the conventional armies from forming ties with locals. By focusing on violence and combat tactics, these studies deviate from earlier explanations for why weak actors are capable of winning irregular wars. Mack (1975) has suggested that small organizations can generate strong cohesion more easily and therefore win against materially superior opponents. Arreguin-Toft (2001) offers a model of strategic interactions between insurgent and incumbent and concludes that weak actors can overcome material inferiority by employing suitable strategies. The common elements in these different approaches to explaining the internal dynamics and outcomes are the actors' abilities to secure civilian loyalties, to mobilize new combatants, and to generate political support. Empiricial studies on the effects of indiscriminate violence have produced contradictory evidence. A series of recent studies have found support for retaliation dynamics that drive civilian loyalties against perpetrators of indiscriminate violence (Condra and Shapiro, 2012, Braithwaite and Johnson, 2012, Linke et al., 2012), but the opposite effect has also been found (Lyall and Wilson, 2009).

The systematic effects of indiscriminate violence remain therefore disputed, both from a theoretical and empirical point of view. Qualitative studies, military accounts, and anecdotal evidence suggest the existence of reactive mobilization as a mechanism in insurgency: Indiscriminate force creates a backlash in public reactions and leads to accelerated mobilization for the opponent. However, this logic is in direct contradiction to the assumption that increasing the risks for the opponent generates incentives for free-riding: Despite heavy losses and dim life expectancies for the individual combatant, some insurgencies have recruited at peak capacity. The Loss Exchange Ratio in Vietnam is usually estimated to be in the magnitude of 1:10 against the insurgents. Palestinian Intifada movements grew strong despite unlikely chances of victory. Afghan Mujaheddin are known for suicide tactics that offer no material incentives to the attacker. Counterinsurgency doctrine, long-range communication, and precision munitions do not seem to have undermined the ability to mobilize in insurgencies around the world. The first open question is therefore the following: *How is it possible for insurgents to mobilize combatants despite heavy losses and dim chances of military success?* Before answering this question, it is important to pay attention to geography as the second recurrent theme in the literature.

The superior ability of irregular forces to travel the desert already informed Lawrence's strategic decisions (Lawrence [1927]1998, 157). Mao ([1938] 1967, 7) assumed guerrilla warfare to be most feasible when employed in large countries (such as China) where the conventional forces of the incumbent or the invader tend to overstretch their lines of supply. Guevara put a higher emphasis on escaping the state's reach through the utilization of difficult terrain (Guevara, 1961, 10).Empirical evidence indicates that rebels indeed seek out the most remote regions to start insurgencies. Macaulay (1978, 288) reports that the Cuban 26th of July Movement operated from the easternmost province of the island – the Sierra Maestra mountains – and then gradually moved towards Havana. Nolan (1958, 71) observes that "in the twentieth century, those seeking power for the purpose of radically transforming society have generally turned to rural-based guerrilla warfare as a means of overthrowing the existing order". Chalk (2008, 5) summarizes the initial phases of the insurgency in southern Thailand with an emphasis on its geographical scope: "the main aims were to present the southern provinces as an area that remained beyond the sovereign control of Bangkok". Salehyan (2009) found that rebels actively utilize international borders to escape the state's reach.

Consequently, counterinsurgency theory has also identified geography as a decisive factor. John Paul Vann was troubled by the low degree of urbanization in South Vietnam in the early 1960s. With an estimated 85% of the population living in the countryside, guerrilla recruitment could take place largely unnoticed by the central government (Sheehan, 1988, 50). Galula (1964, 23) stresses that "the role of geography, a large one in ordinary war, may be overriding in a revolutionary war. If the insurgent, with his initial weakness, cannot get any help from geography, he may well be condemned to failure before he starts". McColl (1969) offers a geographical model that identifies suitable territorial bases for guerrilla movements.

A classic model on the role of distance in interstate wars was presented by Boulding (1962), who assumes that a state's ability to project power toward an enemy is dependent on both its military strength and the distance that separates the adversaries. His notion of a "Loss of Strength Gradient" (LSG) assumes that for every unit of distance, a certain number of personnel have to be subtracted from the fighting forces and added to the supply troops. In theory, putting this relationship into numbers allows us to calculate precisely the geographical limits in power projection. While heavily employed in the context of interstate conflict research (for example Lemke, 1995), the model has also been modified to serve in the context of civil war research (Herbst, 2000, Buhaug and Gates, 2002, Herbst, 2004, Cederman, 2008, Buhaug et al., 2008, Buhaug, 2010). Figure 1 illustrates Boulding's model.

Along these lines, mountainous terrain has been identified as a determinant of war onset. Fearon and Laitin (2003) theorize that especially rough terrain allows a guerrilla movement to escape the state's reach. With larger and more accurate event datasets becoming available to the research community, the question of where fighting takes place in armed conflict has emerged as an independent research topic. The spatial footprint of insurgency has been analyzed for Liberia (Johnston, 2008, Raleigh and Hegre, 2009) and Vietnam (Kalyvas and Kocher, 2009). Moreover, settlement patterns have been linked to both the likelihood of civil conflict as well as the occurrence of conflict events (Weidmann, 2009, Weidmann and Ward, 2010).

With a strong emphasis on geography and civilian population, the existing literature has generated important insights. But while the significance of civilian support and geographic distance in insurgencies has been widely acknowledged, their exact roles remain disputed. Moreover, existing explanations fall short of painting a fully coherent picture.

First, and most importantly, there is an unresolved question of how state reach can explain attributes of modern insurgencies: Boulding's LSG provides an elegant and simple formalism that explains why rebellion is more likely and more successful in a state's periphery. However, the two most prominent advocates of state reach explanations have stressed the predominantly historical relevance of their research:

> The airplane and now the missile have brought about a revolution of quite unprecedented dimensions [...]. For the air-born carrier or weapon, the world is an almost featureless globe: coasts, mountains, deserts, and forests hardly exist as long as there are landing strips on the other side of them. The intricate geographical structure of national power, therefore, which rests on the combination of sea-power with a very low LSG and land-power with a much higher LSG, has largely been swept away as far as air-power is concerned. Everywhere now is accessible to everybody; there are no nooks, corners, or retreats left, and no snugly protected centers of national power (Boulding, 1962, 272).

Similarly, Scott (2009, xii) remarks that his own theory of statehood's inability to permanently integrate the Southeast Asian highlands is largely irrelevant for the post-World War II period. Scott assumes that terrain friction accounts for a large variety of political attributes of the region, including the geographic shapes of states, the spatial distribution of literacy, and much of its early modern history of political violence. However, with the infrastructural and logistical advancements of the last sixty years, terrain friction has lost much of its significance according to Scott.

Empirically we find that rebel victories have become more frequent in the past two centuries despite the overall trends towards more mechanized forces and increased long-range transport capabilities of conventional armies (Arreguin-Toft, 2001, Lyall and Wilson, 2009). It seems therefore that the limited ability to project military force into the periphery is an insufficient explanation for why conventional armies are defeated by irregular insurgencies. Any air force today has the technical ability to project its force within the boundaries of its home state. Moreover, successful insurgencies have been won against superpowers with *global* strike capabilities, such as the United States and the Soviet Union in Vietnam and Afghanistan, respectively. The resulting puzzle is therefore: *How is it possible that irregular insurgencies are successful despite advanced power projection capabilities of conventional armies?*

The next section will attempt to resolve these theoretical questions and derive an explanation that integrates geography, population, and reactive mobilization. After that, two aggregate properties of insurgencies – their military outcomes and approximate levels of casualties – are predicted based on the empirical implications of this theory. This step is not to be confused with a thorough test of all aspects of the theory, but merely a first empirical step toward a comprehensive test.

3 Theoretical Synthesis

Drawing on a distance-decay mechanism largely inspired by Boulding (1962), Kalyvas' (2006) typology of violence, as well as technological and cognitive considerations, I develop a new theoretical model of reactive mobilization in this section. Most importantly, a strong link between geographical conditions, types of applied violence, and the resulting effects for war outcome and casualties are illustrated.

3.1 The Loss of Accuracy Gradient

Boulding's LSG might be suitable to estimate a state's potential reach, but it does not take into account the type of violence that is applied. Kalyvas' typology of violence provides some of the building blocks for a more advanced understanding of the complex dynamics of mobilization that drive casualties and outcomes. His theory focuses on the use of violence as a means of enforcing collaboration and deterring against defection (Kalyvas and Kocher, 2007a, 210). Mobilization and civilian collaboration are largely endogenous to military control in this model (Kalyvas, 2006, 12; 118-132), but the question of how exogenous conditions determine the effect of violence on mobilization cannot be answered solely on this basis.

I therefore propose a synthesis of these models in order to explain outcomes and casualties in irregular civil wars. Several additional assumptions are necessary to make this new theory work. Most importantly, the application of violence is severely restricted by technological and cognitive factors. Let us consider technological aspects first. A nuclear bomb can never be used to apply violence selectively. It is inherently indiscriminate. Similar in this regard but less devastating is conventional, strategic bombing: Explosives in free fall have a greater potential for indiscriminate destruction than direct small arms fire. No one has expressed this logic more clearly than John Paul Vann:

> This is a political war and it calls for discrimination in killing. The best weapon for killing would be a knife, but I'm afraid we can't do it that way. The worst is an airplane. The next worst is artillery. Barring a knife, the best is a rifle — you know who you're killing (Vann, quoted in Sheehan, 1988, 317).

Vann's insight is important because it stresses both the importance of discrimination in the application of violence as well as the technological requirements that are necessary to achieve discrimination. These technological constraints are somewhat symmetrical for incumbent and insurgent. For an insurgent movement, the available forms of violence also range from selective to mostly indiscriminate: from the targeted assassination of individuals, to hit-and-run attacks on strategic targets with the possibility of civilian casualties, to terrorism targeting civilians. Equally symmetrical for both actors is the trade-off between range and accuracy in the application of violence. At close range, both actors can selectively target individuals by drawing on arrests and assassinations. At medium range, both actors risk civilian casualties in their attempts to kill enemy combatants via artillery strikes, tactical bombardments, and hit-and-run attacks. At long range, civilian casualties become inevitable in strategic bombing and international terrorism. Such tactics tend to apply violence indiscriminately and can be applied globally. Individual arrests and targeted assassinations are prototypical examples of selective violence, but their application is mainly confined to the areas under the actors' control. *We can therefore assume that long-range attacks are – on average – less discriminate than short-range attacks and we expect civilian alienation to be a function of distance as well.* Therefore, it is not necessarily the *quantity* of force that diminishes over distance as assumed by Boulding, but its *quality*: Arms branches with a low LSG tend to apply force indiscriminately. In Vann's words, a knife should be the weapon of choice in a political war. But a knife combines the highest accuracy with the shortest range. An airplane combines the lowest accuracy with the highest range. In between the two, there are rifles and artillery that follow the same trade-off.

This insight echoes the classic literature (Arendt, 1970, 53) and provides a new interpretation of Kalyvas' (2006) theory of indiscriminate violence: Instead of assuming that the quality of violence is endogenous to levels of military control, this approach suggests that exogenous geographical constraints to the actors' abilities to identify and selectively target the adversary explain indiscriminate violence. Like Boulding's (1962) theory, this mechanism should not be read as geographic determinism. Actors might very well have different intentions and abilities to apply violence selectively close to their power centers, but the above reasoning suggests that this ability generally declines as power is projected over large distances.

The advancement of arms technology has only marginally contributed to solving this problem. Today, manned and unmanned aerial vehicles are certainly capable of hitting designated targets with great precision. But equating the ability to hit targets reliably with the ability to hit the right target effectively means to confuse precision and accuracy. The dart game analogy can help clarify this distinction: Throwing darts accurately means to loosely group them around the bullseye leading to a low average distance from the intended target. Throwing darts precisely means to tightly group them *somewhere* on the dartboard with similar distances from the intended target. With regard to the application of violence in (counter-)insurgencies, this distinction is crucial: An accurate application of force would reliably affect only enemy combatants while leaving innocent bystanders unharmed. A precise application of force merely amounts to hitting what one is shooting at in combat situations. For example, calling in an air- or artillery strike on a defined location in response to small arms fire will likely result in precisely hitting this location, but it might still harm bystanders. Responding accurately to small arms fire would entail the somewhat unrealistic approach of chasing after the shooter, identifying him beyond reasonable doubt, then taking appropriate action according to the criminal law of the state or the Geneva Convention. Clearly, technological advancements along the lines of guided munitions have enabled violence to be applied with greater precision, but not with greater accuracy.

Improving the latter seems especially difficult since cognitive accuracy also declines as a function of distance. As long as language and customs are known, the identification of loyalties is comparatively easy. Disagreement with the government can be inferred from subtle cues. Subtones and connotations can give away collaborators. Clothing, appearance, and dialect allow inferences regarding ethnic and religious ties. Knowledge of local history allows for a better understanding of local grievances. These insights are part of modern counterinsurgency doctrine:

> Learn about the people, topography, economy, history, religion, and culture of the area of operations (AOs). Know every village, road, field, population group, tribal leader and ancient grievance. Become the expert on these topics (DOD, 2007, Appendix A, Paragraph A2).

This demand is reasonable in theory, but increasingly difficult to achieve in practice as distance to the actor's cultural circle grows. Europeans and Americans in Algeria and Vietnam found a variety of ethnic, linguistic, and cultural aspects that were mind-boggling to the outsider and hard to generalize from. Many historians and sociologists have struggled to paint a coherent picture of Afghan society and the most successful attempts require an introduction on social units and concepts that are foreign to most Westerners (Dorronsoro, 2005). Mental access to foreign regions is hard to acquire and hard to communicate.

Especially for foreign troops in counterinsurgency campaigns, the impression of being in pointless pursuit of an invisible enemy is commonplace: "The Vietcong were so intermingled with the peasantry that the Saigon troops had difficulty distinguishing friend from foe. [...] How much more difficult it would be for Americans. The American soldier would soon start to see the entire rural population as the enemy [...]"(Sheehan 1988, 383; see also: Ellsberg 2003, 167). Arguably, this problem becomes more severe as troops leave government-controlled areas and advance into the periphery. The absence of cognitive access to local customs can provide incentives for brutal and indiscriminate behavior. Third-party actors in counterinsurgency operations are certainly more prone to the problem of not understanding foreign societies. But even governments fighting insurgencies within their own territorial boundaries face this problem. While insurgents often mobilize local combatants for local operations, draft-based armies usually employ soldiers in areas unknown to them.

One way to circumvent the problem of limited cognitive access to remote regions is the resettlement of the civilian population. This approach has been tried as part of the Briggs Plan by the British in Malaya (Hack, 2009, 4) and in the context of the Strategic Hamlet Program by the US in Vietnam (Sheehan, 1988, 309). The success of such measures remains questionable at best: The ability to apply violence more selectively in and around newly established hamlets does not necessarily lead to less alienation because the inhabitants of these hamlets must be forcefully relocated. Therefore, the advantage arising from increased accuracy can be nullified by the alienation that results from forced relocation and destruction of property.

Several implications arise from this discussion. First, unlike Boulding's model, the presented theory assumes that the quality (not primarily the quantity) of force is a function of distance in modern warfare. It extends his well-established “Loss of Strength Gradient” into a “Loss of Accuracy Gradient”. As the distance between incumbent and insurgent grows, the accuracy of their violent action decreases. Compensating for the lack in accuracy with increased lethality only amplifies the problem: More innocent bystanders fall victim to the inaccurate use of force. Second, indiscriminate force usually creates a backlash in public opinion. The alienating effect of inaccurate violence vastly outweighs its deterrent effect, since even supporters of the user of force can fall victim to its measures. This logic is illustrated in Figure 2.

Negative consequences of indiscriminate violence for the perpetrator have been proposed theoretically for a long time (Ellsberg, 1970, Bennett, 2008, Kalyvas, 2006, 144;151). More recently, empirical support for this claim has been found (Downes, 2007, Braithwaite and Johnson, 2012, Condra and Shapiro, 2012, Linke et al., 2012), although some studies have produced opposite results (Lyall, 2009).

Third, one would expect theaters of war that see mainly indiscriminate violence to give rise to fiercer resistance and counter-resistance. In these scenarios, reactive violence repeatedly leads to mobilization and counter-mobilization and fuels the conflict over multiple iterations. If one actor enjoys superior accuracy due to its proximity to the bulk of the population, it will win the civil war. This effect plays out in a somewhat counterintuitive manner since the winning side might very well suffer heavier losses. However, a constant stream of newly alienated civilians can fill the ranks of the fallen and make a prolonged campaign sustainable.

Accounting for the systematic effect of declining accuracy and rising resentment offers a theoretical solution to the unresolved questions mentioned above. This theory rejects limited state reach and instead proposes limited accuracy in the application of force as an explanation for why insurgencies succeed. It also rejects the idea of material self-interest as a fully sufficient explanation for why civilians take sides in insurgencies. Victims of violence and their kin might simply seek revenge and strive to fight injustice. Translated into the terminology of the collective action problem, inaccurate force provides revenge as a selective incentive to victims and their kin. Moreover, accurate and conditional violence close to the actors' centers should have a stronger coercive effect than largely inaccurate and unconditional violence projected over long distances. This mechanism is also symmetric for both insurgent and incumbent. Arbitrary and unjustified violence are often attributed to the state in national liberation movements. Similarly, states can increase their domestic support for counterinsurgencies by pointing to indiscriminate violence used by the rebels in the form of terrorism.

Having derived the Loss of Accuracy Gradient from both Kalyvas and Boulding, I derive empirical im-

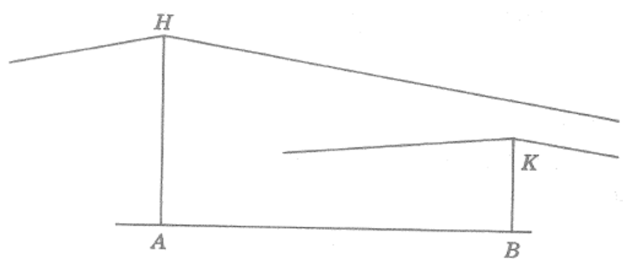

Figure 1: Boulding's original LSG concept: states A and B have unequal military powers. Although A's power declines over distance, it is still superior to B's even in B's capital.

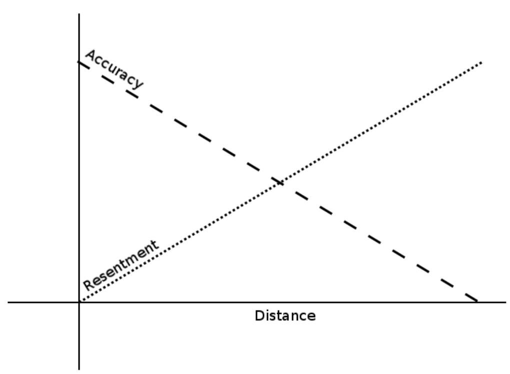

Figure 2: The "Loss of Accuracy Gradient" is illustrated in this figure. As distance to the power center increases, accuracy declines. This leads to an increase in resentment toward the user of force.

plications in the next section. It must be mentioned that this analysis necessarily falls short of empirically proving all aspects of the presented theory as it is beyond the possibilities of a single paper. Instead, a cross-sectional analysis of two macro-level properties of insurgencies – their military outcomes and the overall levels of battle-related fatalities – are tested and shown to be in line with the theoretical expectations.

3.2 Hypotheses

The two main hypotheses directly result from the theory: The actor that is able to exercise violence more accurately by being closer to the bulk of the population is more likely to win the civil war. Moreover, whenever both actors exercise inaccurate violence, the war will be more severe than if one side enjoys superiority. Operationalizing accuracy can be achieved based on the assumption of a Loss of Accuracy Gradient: With growing distance from the actors' power centers the quality of their applied violence deteriorates, as depicted in Figure 2.

Mao's "Protracted War" model explicitly assumes that fighting in the initial stages of an insurgency takes place in the periphery. Once the more remote regions are secured, rebel forces advance toward the capital city in multiple stages of their campaign. The communist insurgencies in China, Cuba, and Vietnam, as well as the anti-Soviet insurgency in Afghanistan, all reflect this pattern. This allows for the simplifying assumption that the center of state power can be associated with the capital city and that the rebels' realm is the periphery. This assumption is in line with the communist literature (Guevara, 1961, 10), counterinsurgency studies (Galula, 1964, 23-24), recent conflict research (Kalyvas and Balcells, 2010, 415), and agent-based simulation studies (Cederman, 2008).

A suitable proxy for the government's ability to control the civilian population with accurate force is therefore the capital's distance to the bulk of the population. Similarly, the rebels' ability to apply force accurately over large fractions of the population relies on civilians being concentrated in the periphery. Summing up these considerations, the following hypothesis results from the theory:

- H1: Distance between population concentrations and the capital has a positive effect on insurgent victory.

The second main hypothesis expresses the other central implication of the theory: The territorial balance that determines the actors' ability to apply force accurately affects the number of casualties in insurgencies: If one side enjoys far superior accuracy due to proximity to the bulk of the population, the indiscriminate tactics of its opponent quickly undermine civilian support, leading to a comparatively bloodless struggle. However, if the bulk of the civilian population is located at medium distances to both power centers, it suffers from indiscriminate attacks by both sides.

The resulting reactive mobilization feeds into the circle of violence since it only results in the application of indiscriminate force. Multiple iterations of violence, alienation, and mobilization generate higher levels of deaths and destruction. The main hypothesis regarding casualties therefore expresses this inverse U-shaped relationship:

- H2: Distance between population concentrations and the capital has a negative quadratic effect on the number of war casualties.

Again, one must control for other variables affecting access to the population. Urban population and land-cover are therefore taken into account in the empirical models. The following section presents empirical tests of these hypotheses followed by a general discussion.

4 Empirical Analysis

Drawing on a dataset by Lyall and Wilson (2009) that provides information on the outcome of insurgencies as well as data on casualties from the Correlates of War Project (Sarkees and Wayman, 2010) and the CSCW/PRIO battle death data (Lacina and Gleditsch, 2005a), the proposed theory is tested in a large-N conflict-level setting. Investigating the macro-effects of the proposed theory was chosen as a suitable first step, since a detailed test of the different aspects of the theory would be beyond the capabilities of a single paper. Moreover, studies on lower levels of aggregation would necessarily fall short of testing the proposed macro connection of geography, outcome, and casualties. For example, case studies with conflict events as the unit of analysis could shed light on single aspects of the theory, such as reactive mobilization or the distance decay of violence, but they would fall short of explaining aggregate properties such as outcomes and casualties. An analysis on the level of ethnic groups would be possible in principle, but casualty data is not available on this level and the analysis would have to be restricted to ethnic insurgencies.

Based on these considerations, I decided to operationalize the actors' proximity to the bulk of the population for the country level, making this study compatible with other quantitative studies of insurgency (Fearon and Laitin, 2003, Lyall and Wilson, 2009, Connable and Libicki, 2010). Building directly on a dataset by Lyall and Wilson (2009) enables the problem to be researched in a cumulative fashion by explicitly highlighting how new insights contribute to the explanation of political processes beyond established knowledge (see: Hegre and Sambanis, 2006). However, certain limitations also arise from a conflict-level analysis and these should be briefly discussed. While the theoretical discussion focuses on explaining the dynamics of insurgent mobilization and violence, a conflict-level research design is essentially static. The empirical test therefore only indicates that the initial conditions of the conflict and its eventual outcome are in line with the proposed theory without testing the exact causal pathway. As such, the empirical analysis only constitutes a first step towards testing the theoretical model. A conflict-level measurement, the *Territorial Balance Indicator* (TBI), will be introduced below, followed by the empirical results and a discussion of their implications.

4.1 Measuring Territorial Balance

The proposed theory puts great emphasis on distance as a crucial determinant of accuracy in the application of violence. Distance between the bulk of the population and the military actors therefore must be defined to test for these proposed effects. Data on the spatial distribution of population –the "Gridded Population of the World Dataset" (GPW) (see: CIESIN, 2005)– and the location of capital cities and international boundaries (Weidmann et al., 2010) were combined to form a single indicator for "Territorial Balance". Intuitively speaking, Territorial Balance measures the fraction of the total population within a certain distance from the capital city. For any given distance D from the capital, the fraction of the population $p_{d<D}$within that distance can be calculated and plotted, according to $tb_D = \frac{\sum p_{d<D}}{\sum p}$,where p is the total population of a country and d the distance from the capital to the most remote inhabited spot. This measurement naturally falls into the interval $[0, 1]$. With growing distance, tb_D values form a curve eventually connecting the points (0,0) and (1,1). Figure 3 shows the resulting curve for the United States, based on the 1990 spatial distribution of the population.

How can we interpret the exact shape of the curve in Figure 3? The curve simply expresses the fraction of the population that can be reached for any given distance from Washington DC. At first, the curve climbs rapidly as the densely populated east coast is covered. Around a cumulative population value of 0.4, the curve flattens out as the less densely populated Midwest is reached. It flattens out even further as the Rocky Mountains are integrated and then suddenly spikes around a distance value of 0.7 caused by the densely populated west coast. The sparsely populated wilderness of Alaska and the vast Pacific Ocean between Hawaii and the continent appear as an almost flat section on the right end of the curve.

For the statistical analysis, the described indicator must be collapsed into a single number for which the area under the curve (AUC) was chosen. Formally: $TBI = \frac{a}{a+b}$, where a is the area under the curve and b the area above it. A caveat of this approach is that global data on population numbers only dates back to 1990. Therefore, the post-WWII period that is frequently the focus of civil war research cannot be analyzed entirely based on this indicator.[2]

4.2 Analyzing Outcome

The possible outcomes in insurgencies can be categorized on a spectrum ranging from victory to conclusive defeat for either side. To preserve the ordinal information in the dependent variable, I coded incumbent victory as 1, settlement of the hostilities without decisive victory of either side as 2, and insurgent victory as 3, and Ordered Logistic regression models were used to assess the relevance of the proposed variables. The time period under investigation is 1970 to 2010. The reason for limiting the time period to this range lies in the static nature of the main independent variable. The TBI in its current form relies on geo-referenced data that can only be obtained from 1990 onwards. Since the aggregate TBI measure is not subject to drastic annual changes, 1990± 20 years was chosen as a suitable time frame for the analysis which uses static 1990 TBI values as the main independent variable.[3] For countries whose international boundaries have changed over time, I calculated the TBI for the boundaries that were in place when the conflict broke out.

[2] This measurement is of course not the only possible operationalization of access to the civilian population for the military actors. As a robustness check, I have coded an alternative version of the indicator using distances to the nearest major city. The cities coded in this dataset (Nelson, 2008) had at least 50,000 inhabitants in the year 2000, and most of them existed back in 1970. The results are reported in the Appendix and are substantively identical to the ones presented in this paper.

[3] To rule out endogeneity, i.e. systematic effects of insurgencies on the TBI, I compared TBIs calculated based on 1990 and 2010 GPW data for conflict and non-conflict cases. For countries that experienced conflict between 1990 and 2010, the average change in TBI was -0.0014, while non-conflict countries changed with a slight increase to 0.0041. These changes are roughly equivalent to 0.1 to 0.2 standard deviations of the entire 1990 TBI sample. In both cases, no systematic differences in the TBI distributions were found in Kolmogorov-Smirnov tests. Visual comparisons of the TBI distributions and side-by-side comparisons of the TB curves for conflict countries can be found in the Appendix. In order to rule out multicollinearity in the data, I calculated correlations between the main independent variables. The resulting table can be found in the Appendix and shows no signs of strong multicollinearity.

Mounting methodological objections against a sole reliance on p-values (Ward et al., 2010, Schrodt, 2010) require additional assessments of the predictive capabilities of newly introduced variables. Such tests directly communicate the external validity of statistical models.

Both in-sample and out-of-sample predictive performance are included in the empirical analysis. To create a baseline for the statistical analysis, model 1 uses the model from the Lyall and Wilson (2009) study as a baseline. The main independent variable of their study – modern combat tactics – had to be omitted in this analysis. Modern combat tactics according to Biddle's (2006) definition were generally introduced after WWI. Therefore, there is simply no variance on this indicator for post-1970 insurgencies. However, Lyall and Wilson (2009) still provide a formidable point of departure both theoretically and empirically.

Model 2 only uses distance from the capital city of the incumbent to the conflict country and the TBI. The distance variable is equal to 1 in all cases where incumbents fight domestic insurgencies, but varies for out-of-area conflicts. Model 3 includes all introduced variables. Model 4 omits the distance variable and shows that the results do not change substantively in comparison to the full model. Model 5 was chosen to minimize the AIC value, a statistic that rewards high goodness-of-fit while penalizing for the number of explanatory variables.[4]

4.2.1 Control Variables

Military and socioeconomic control variables were obtained from Lyall and Wilson (2009) and supplemented with the main independent variable. As visible at first glance in Table 4, several control variables are significantly associated with the ordinal coding of rebel success.

Regime type one year prior to conflict onset (*REGIME*) and outside support to the rebels during the conflict (*SUPPORT*) are significantly and positively associated with rebel victory.[5] The positive sign for the *REGIME* variable seems to confirm the widespread suspicion that democratic governments are likely to give up fighting counterinsurgencies earlier than autocracies. Domestic peace movements, for example, can put significant pressure on the incumbent in liberal and democratic societies. Autocratic systems, on the other

[4]The central result of the analysis holds for a wide range of model specifications. However, it was found that even the base model for the post-1970 period leads to quasi-complete separation: Extremely high values of the DISTANCE are consistently associated with rebel victory. Omitting the DISTANCE variable leads to essentially the same estimates for the remaining covariates, as model 4 shows.

[5]REGIME was measured by the polity2 variable in the Polity IV dateset (Lyall and Wilson, 2009, Jaggers and Gurr, 1995).

hand, can keep fighting without having to take into account the interests of a broad electorate. Therefore, autocratic societies are more likely to defeat insurgent forces militarily.

It should be mentioned, however, that this finding seems to be driven by the restricted sample size and does not hold for the entire post-WWII period, as reported by Lyall and Wilson (2009). Foreign *SUPPORT* to the rebels is also reliably and positively associated with rebel victory. Foreign fighters willing to join the uprising, material support, and cooperating intelligence services can obviously strengthen the rebel forces. The state capability indicator (*POWER*)[6] is negatively and significantly associated with insurgent victory which again confirms the intuition that stronger states are better at fighting off domestic challengers. The logged energy use of the incumbent as a share of the energy use of the whole population (*ENERGY*) also has a negative estimate, but fails to reach statistical significance. Despite the limited sample and the ordinal coding of the dependent variable, these results are largely in line with the results reported by Lyall and Wilson (2009). An indicator of whether a country is occupied (*OCCUPATION*) is significant in models 1 and 4, but not in the remaining models. Mean elevation of the country (*ELEVATION*) does not reach significance at the 5% level.

Clearly, elevation averaged over the whole country is a rather crude measurement and one should not read too much into this non-effect. The role of inaccessible terrain and limited access to peripheral regions is very central in the theoretical and qualitative literature and probably not adequately operationalized in this variable. A reliable indicator of war outcome is *COLDWAR*. In all four models, *COLDWAR* is negatively associated with insurgent victory, implying that the probability of insurgent success was significantly lower during the Cold War. African countries that received military aid from the superpowers might drive this result: The Democratic Republic of the Congo, Liberia, and Rwanda have all witnessed successful uprisings since the end of the Cold War. The variables *MECH* and *HELI* reflect military capabilities of the incumbent. *MECH* encodes the degree of mechanization of the armed forces and has no significant effect, which deviates from the findings reported in Lyall and Wilson (2009) for the entire post-WWII period. The availability of helicopters (*HELI*) has a positive effect on insurgent victory in all models. The number of languages

[6]POWER was coded as the natural log of the cumulative national capabilities of the incumbent in the last prewar year based on the Correlates of War dataset (v.3.02) (Lyall and Wilson, 2009, Singer and Small, 1994). It is important to mention that the variables *REGIME*, *POWER*, and *ENERGY* reflect the country's situation in the last prewar year, while *SUPPORT*, *HELI*, and *COLDWAR* code whether external rebel support, helicopter deployment, or parts of the Cold War took place at some point during the conflict.

spoken in the conflict country ($LANG$) has a positive estimate in all models, but is only weakly significant in models 4 and 5. This positive estimate is in line with the discussion on cognitive access to the civilian population: Incumbents should face a more profound problem of identifying civilian loyalties if multiple languages are spoken in the conflict-torn country. International $TRADE$ as a fraction of the overall GDP is weakly negatively associated with outcome, as is the natural log of the total $POPULATION$ in the country. Distance from the counterinsurgent's home country to the conflict country ($DISTANCE$) has a small but significant positive effect, which also corresponds to the theoretical expectations.

4.2.2 Main Independent Variable

For all cases of insurgency including expeditionary wars, the TBI was calculated for the country in which the actual conflict took place. The territorial balance indicator (TBI) is strongly and robustly associated with rebel victory. As explained above, TBI values close to 1 indicate that the state has the entire population in the vicinity of the capital city. Values close to 0 correspond to situations in which the state has to reach out to its most remote corners to reach the bulk of the population. Therefore, the large, negative estimate and the small standard error support the theoretical expectation and Hypothesis 1. The TBI remains highly significant with the expected negative sign for all fitted models.

4.2.3 Predictive Performance

The regression models were used to predict probabilities for each possible outcome of a given war. Taking the maximum of these probabilities as the model's prediction allows for a comparison with the empirical record. The absolute value of the mismatch between prediction and empirical record provides a simple distance metric that still takes into account the ordinal scale of the dependent variable. This simple deviation can be formalized as $\frac{\sum |Y-\widehat{Y}|}{N}$ and corresponding results are shown in Table 4. Although it is robustly associated with outcome, the TBI does not improve the predictive capabilities of the models in-sample or out-of-sample. For the out-of-sample tests, leave-one-out cross-validation was performed: The model was fitted on N-1 observations and then used to predict the dependent variable of the missing observation. Again, ordinal deviation was chosen as the measure of success and the process was repeated for all observations. As one can see in Table 4, the baseline model 1 performs quite well in-sample, with an average deviation score of 0.4462.

Only model 5 that includes the TBI is slightly better, while the remaining models are worse. Surprisingly, the minimal model 2 performs best out-of-sample, followed by model 5. In-sample and out-of-sample, the TBI improves statistical predictions. Based on model 4, the effect of the TBI variable on insurgent victory is visible in Figure 5. TBI values observed in the empirical sample have an effect on the predicted probabilities for insurgent victory, but the corresponding error bands are comparatively large for low-TBI cases. For high-TBI cases, the error bands are much smaller.

4.3 Analyzing Casualties

The second main implication of the proposed theory touches upon war severity measured in battle-related casualties. The proposed theory and the derived geographic indicator are to some extent capable of explaining why insurgencies in certain countries are much more lethal than others. Generally, regression models based on the Poisson and Negative Binomial distributions are suitable to model dependent count variables. Based on the results of a likelihood-ratio test, a Negative Binomial model was preferred over a Poisson-based approach. The Negative Binomial model is also better suited for modeling the heavy-tailed casualty distribution than a Poisson model.

The casualty data provided by the Correlates of War project and CSCW/PRIO do not cover the entire sample of insurgencies. Reliable numbers on casualties from these two sources could only be obtained for 57 out of the 64 cases under investigation. The coding procedure used Correlates of War data (Singer and Small, 1994, Sarkees, 2000) wherever available (Version 4.1, 24 cases), and the best casualty estimates in the CSCW/PRIO data (Lacina and Gleditsch, 2005b) to supplement the dataset (Version 3.0, 33 additional cases).

Lacina and Gleditsch (2005b) discuss some of the conceptual and empirical challenges that arise in the systematic analysis of casualty figures in war. I decided to analyze battle-related casualties for testing the presented theoretical argument, thereby excluding those fatalities that were caused by disease, hunger, and other indirect consequences of fighting. The margin of error for battle deaths data is comparatively high and the data can be biased, which follows from the limited ability to cross-validate casualty claims and the impossibility of accounting for all casualties in the turmoil of war (see: Oliver and Myers, 1999, Davenport and Ball, 1996, Kalyvas, 2004, Siegler et al., 2008). Other recent studies have found that media-based conflict-event datasets code a representative sample of the conflict events obtained from military records (O'Loughlin et al., 2010). With the reliability of casualty data in question, I tried to use it conservatively and excluded cases where insurgencies took place in the context of wider interstate wars and also dropped observations where no best estimate for at least one of the conflict years was available.

Model 6 in Table 4 on page 33 is used again as a baseline model for explaining casualties. Model 7 is a minimal model containing only the two functional forms of the TBI. Model 8 shows results for the full model. Model 9 was chosen to minimize the AIC statistic. In all four models, the dependent variable is the number of battle-related fatalities for each war.

4.3.1 Control Variables

As visible in the regression table 4, several control variables are significantly associated with battle-related casualties. The *REGIME* variable is negatively associated with casualties (weakly significant), indicating that more democratic regimes are likely to be involved in less severe counterinsurgency wars. External *SUPPORT* is robustly associated with higher casualty figures, which corresponds to the intuition that fueling the conflict from the outside leads to more casualties. *POWER*, reflecting the state's military capabilities, is not robustly related to casualties, but is generally estimated to be positive. The *ENERGY* variable has a negative estimate, but statistical significance is not reached. *OCCUPATION* is estimated to be negative and is found to be significant in all models. Possibly, insurgencies against occupying forces leave the incumbent with a natural "exit strategy" in terms of allowing for a retreat from the theater. Domestic incumbents might not have this option, which leads to more devastating wars within the sample.

ELEVATION and *DISTANCE* do not seem to have a large effect on casualties. The estimates for *COLD-WAR* are positive and highly significant. The Cold War witnessed committed support to both insurgents and incumbents by the superpowers, resulting in a series of intense conflicts such as the struggle for South Vietnam and the Afghan civil war following the Soviet invasion. *MECH* – the degree of mechanization of the armed forces – generally has a positive effect on casualties. As Lyall and Wilson (2009) point out, highly mechanized armies are less capable of forming ties with the civilian population. This means that their ability to tell friend from foe is less developed and the probability of applying inaccurate violence is higher. The combination of low accuracy in the application of force combined with the high lethality of modern arms might be driving this result. The *HELI* variable does not seem to have a significant effect on casualties.

4.3.2 Main Independent Variable

The quadratic TBI specification is robustly highly significant with a large negative estimate. The inverse U-shaped relationship between TBI and the number of fatalities in insurgencies can therefore be found in the empirical record, as seen in Figure 6. This finding lends support to Hypothesis 2.

4.3.3 Predictive Performance

With regard to casualties, the predictive performance of models 6 to 9 was also evaluated. Average deviation between predicted and observed casualties was used as a criterion. Cross-validation was again performed for out-of-sample testing. As shown in Table 4, the TBI indicator strongly improves casualty predictions. In direct comparison, base model 6 performs worse than the TBI-only model 7. In-sample, the absolute average deviation in predicted casualties is 855.2186 for model 6, but only 828.004 for model 7. Out-of-sample, model 6 mispredicts the number of casualties on average with 1430.658 and model 7 with only 860.0593. The lowest AIC is achieved with model 9 which includes the quadratic form of the TBI.

5 Discussion and Conclusion

This study has introduced a new synthesis between Boulding's (1962) classic distance-decay model and Kalyvas' (2006) typology of violence. Focusing on the distance-decay of the *quality* of violence and reactive mobilization, the study has identified the spatial distribution of population as a decisive factor in insurgencies that predicts both military outcomes and approximate levels of casualties.

These results corroborate other recent empirical insights suggesting that the role of military power in counterinsurgencies is counterintuitive and less decisive than in interstate wars. A corresponding quantitative indicator has been proposed and tested in a conflict-level analysis. This analysis supports the theoretical expectation (Hypothesis 1) that the military actors' access to the population is associated with the outcome of insurgencies and improves out-of-sample predictions.

Apart from this indicator, institutional and military variables are also robustly related to outcomes in insurgencies, as Lyall and Wilson (2009) have shown. The results therefore should not be read as a geographic determinism, but a robust probabilistic tendency in line with the theoretical expectations.

With regard to casualties (Hypothesis 2), the proposed theory has passed another empirical test. The characteristic, inverse U-shaped relationship between territorial balance and casualties can be found in the empirical record, which corresponds to the expectation that civilians are alienated and reactively mobilized by both sides if the bulk of the population resides at medium distances from the capital city. In this second analysis, the quadratic TBI variable also strongly improves predictions, both in- and out-of-sample. An alternative specification of the TBI using distances to major cities yields substantively identical results as discussed in the Appendix. These results are very encouraging, but they are merely a first empirical step and necessarily fall short of proving all aspects of the theory. In order to gain more empirical traction, conflict-event data from recent insurgencies can be used to test for the proposed reactive patterns and the link between the type of violence and location.

At lower levels of aggregation, fine-grained data on reactions to violence in insurgencies, the spatial extent of insurgencies, and the compositions of uprisings and counterinsurgencies in terms of local versus foreign combatants could shed more light on the empirical reality of the proposed mechanisms.

In summary, an integrated theory has been presented that accounts for the causes and effects of indiscriminate violence in insurgencies. This theory explains why increased power projection capabilities alone are insufficient to guarantee incumbent victory and suggests that rebels can be motivated to retaliate instead of being deterred by indiscriminate violence. Since constraints on the application of selective violence in civil wars are based on largely static geographical conditions, outcomes and casualties could be predicted ahead of time in principle. A prerequisite for relying on such predictions would be an increase of their reliability. Future research will have to show to what extent the proposed theory correctly predicts effects on lower levels of aggregation and to what extent its predictive powers on the conflict level can be improved.

References

Allen, J. (2003). *Lost Geographies of Power.* Backwell Publishing.

Arendt, H. (1970). *On Violence.* HMH Books.

Arreguin-Toft, I. (2001). How the weak win wars: A theory of asymmetric conflict. *International Security 26*(1), 93–128.

Bennett, S. D. (2008). Governments, civilians, and the evolution of insurgency: Modeling the early dynamics of insurgencies. *Journal of Artificial Societies and Social Simulation 11*(47).

Berman, E., M. Callen, J. H. Felter, and N. Shapiro, Jacob (2011). Do working men rebel? insurgency and unemployment in afghanistan, iraq, and the philippines. *Journal of Conflict Resolution 55*(4), 496–528.

Biddle, S. (2006). *Military Power*. Princeton University Press.

Boulding, K. (1962). *Conflict and Defense: A General Theory*. Harper.

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

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

Buhaug, H., L.-E. Cederman, and J.-K. Rød (2008). Disaggregating ethno-nationalist civil wars: A dyadic test of exclusion theory. *International Organization 62*(3), 531–551.

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

Buhaug, H. and K. S. Gleditsch (2008). Contagion or confusion? why conflicts cluster in space. *International Studies Quarterly 52*, 215233.

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

Chalk, P. (2008). The malay-muslim insurgency in southern thailand: Understanding the conflict's evolving dynamic. RAND Counterinsurgency study, Paper 5.

CIESIN (2005). Gridded population of the world, version 3 (gpwv3). Columbia University; and Centro Internacional de Agricultura Tropical (CIAT). 2005. Gridded Population of the World Version 3 (GPWv3):

Population Grids. Palisades, NY: Socioeconomic Data and Applications Center (SEDAC), Columbia University. Available at http://sedac.ciesin.columbia.edu/gpw.

Collier, P. (2000). Rebellion as a quasi-criminal activity. *Journal of Conflict Resolution 44*(6), 839–853.

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

Connable, B. and M. C. Libicki (2010). *How Insurgencies End*. National Defense Research Insitute.

Davenport, C. and P. Ball (1996). Views to a kill: exploring the implications of source selection in the case of guatemalan state terror, 1977-1996. *Journal of Conflict Resolution 46*(3), 427–50.

DOD (2007). *U.S. Army Counterinsurgency Handbook*. Skyhorse Publishing.

Dorronsoro, G. (2005). *Revolution Unending: Afghanistan, 1979 to the Present*. Columbia University Press.

Downes, A. B. (2007). Draining the sea by filling the graves: Investigating the effectiveeness of indiscriminate violence as a counterinsurgency strategy. *Civil Wars 9*(4), 420–44.

Ellsberg, D. (1970). Revolutionary judo. Available online at http://www.rand.org/pubs/documents/2006/D19807.pdf.

Ellsberg, D. (2003). *Secrets: A Memoir of Vietnam and the Pentagon Papers*. Penguin Press.

Fall, B. B. (1965). The theory and practice of insurgency and counterinsurgency. *Naval War College Review*.

Fearon, J. D. and D. D. Laitin (2003). Ethnicity, insurgency and civil war. *American Political Science Review 97*(1), 75–90.

Galula, D. (1964). *Counterinsurgency Warfare: Theory and Practice*. Praeger Security International.

Greiner, B. (2009). *War without Fronts – The USA in Vietnam*. The Bodely Head.

Grossman, H. (1999). Kleptocracy and revolutions. *Oxford Economic Papers 51*(2), 267–83.

Guevara, E. (1961). *Guerrilla Warfare: Authorized Edition*. Ocean Press.

Hack, K. (2009). The malayan emergency as counter-insurgency paradigm. *Journal of Strategic Studies 32*(3), 383–414.

Hastings, M. (2010). The runaway general. *Rolling Stone Magazine* (RS 1108/1109).

Hegre, H. and N. Sambanis (2006). Sensitivity analysis of empirical results on civil war onset. *Journal of Conflict Resolution 50*(4), 508–535.

Herbst, J. (2000). *States and Power in Africa - Comparative Lessons in Authority and Control*. Princeton University Press.

Herbst, J. (2004). African militaries and rebellion: The political economy of thread and combat effectiveness. *Journal of Peace Research 41*(3), 357–369.

Jaggers, K. and T. R. Gurr (1995). Transitions to democracy: Tracking democracy's third wave with the polity iii data. *Journal of Peace Research 32*(11), 469–82.

Johnston, P. B. (2008). The geography of insurgent organization and its consequences for civil wars: Evidence from liberia and sierra leone. *Security Studies 17*, 107–137.

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

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

Kalyvas, S. (2006). *The Logic of Violence in Civil Wars*. Cambridge University Press.

Kalyvas, S. and L. Balcells (2010). International system and technology of rebellion: How the end of the cold war shaped internal conflict. *American Political Science Review 104*(3), 415–429.

Kalyvas, S. and A. Kocher, Matthew (2007a). Ethnic cleavages and irregular war: Iraq and vietnam. *Politics and Society 35*(2), 183–223.

Kalyvas, S. and M. A. Kocher (2007b). How free is "free riding" in civil wars? violence, insurgency, and the collective action problem. *World Politics 59*(2), 177–216.

Kalyvas, S. and M. A. Kocher (2009). The dynamics of violence in the vietnam war: An analysis of the hamlet evaluation system (hes). *Journal of Peace Research 46*(3), 335–355.

Kilcullen, D. (2009). *The Accidental Guerrilla: Fighting Small Wars in the Midst of a Big One*. Oxford University Press.

Lacina, B. and N. P. Gleditsch (2005a). Monitoring trends in global combat: A new dataset of battle deaths. *European Journal of Population 21*(2-3), 145–166.

Lacina, B. and N. P. Gleditsch (2005b). Monitoring trends in global combat: A new dataset of battle deaths. *European Journal of Population 21*(2-3), 145–166.

Lawrence, T. E. (1998). *Revolt in the Desert*. Combined Publishing.

Lemke, D. (1995). The tyranny of distance: Redefining relevant dyads. *International Interactions 21*(1), 23–38.

Lichback, M. I. (1995). *The Rebel's Dilemma*. University of Michigan Press.

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

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

Lyall, J. and I. Wilson (2009). Rage against the machines: Explaining outcomes in counterinsurgency wars. *International Organization 63*(1), 67–106.

Macaulay, N. (1978). The cuban rebel army: A numerical survey. *The Hispanic American Historical Review 58*(2).

Mack, A. (1975). Why big nations lose small wars: The politics of asymmetric conflict. *World Politics 27*(2), 175–200.

Mao, T.-t. (1961). *On Guerrilla Warfare*. University Of Illinois.

Mao, T.-t. (1967). *On Protracted War.* Foreign Language Press.

Mason, D. T. and D. A. Krane (1989). The political economy of death squads: Toward a theory of the impact of state-sanctioned terror. *International Studies Quarterly 33*(2), 175–198.

McColl, R. (1969). The insurgent state: Territorial bases of revolution. *Annals of the Association of American Geographers 59*(4), 613–631.

Nelson, A. (2008). Estimated travel time to the nearest city of 50,000 or more people in year 2000. Available online at http://bioval.jrc.ec.europa.eu/products/gam/.

Nolan, D. (1958). From foco to insurrection: Sandinista strategies of revolution. *37*(4), 71–84.

Oliver, P. E. and D. J. Myers (1999). How events enter the public sphere: Conflict, location, and sponsorship in local newspaper coverage of public events. *American Journal of Sociology 105*, 38–87.

O'Loughlin, J. and F. Witmer (2010). The localized geographies of violence in the north caucasus of russia, 1999-2007. *Annals. Association of American Geographers 100*(3), 2379–2396.

O'Loughlin, J., F. Witmer, A. Linke, and N. Thorwardson (2010). Peering into the fog of war: The geography of the wikileaks afghanistan war logs 2004-2009. *Eurasian Geography and Economics 51*(4), 472–95.

Olsen, M. (1965). *The Logic of Collective Action: Public Goods and the Theory of Groups.* Harvard Economic Studies.

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

Salehyan, I. (2009). *Rebels without Borders.* Cornell University Press.

Sarkees, M. R. (2000). The correlates of war data on war: An update to 1997. *Conflict Management and Peace Science 18*(1), 123–144.

Sarkees, M. R. and F. Wayman (2010). *Resort to War: 1816 - 2007.* CQ Press.

Schrodt, P. A. (2010). Seven deadly sins of contemporary quantitative political analysis. Paper prepared for the theme panel 'A Sea Change in Political Methodology?' at the Annual Meeting of the American Political Science Association, Washington, 2 - 5 September 2010.

Scott, J. C. (2009). *The Art of Not being Governed*. Yale University Press.

Sheehan, N. (1988). *A Bright Shining Lie: John Paul Vann and the USA in Vietnam*. Jonathan Cape.

Siegler, A., L. Roberts, E. Balch, E. Bargues, A. Bhalla, C. Bills, E. Dzeng, Y. Epelboym, T. Foster, L. Fulton, M. Gallagher, J. D. Gastolomendo, G. Giorgi, S. Habtehans, J. Kim, B. McGee, A. McMahan, S. Riese, R. Santamaria-Schwartz, F. Walsh, J. Wahlstrom, and J. Wedeles (2008). Media coverage of violent deaths in iraq: An opportunistic capture-recapture assessment. *Prehospital and Disaster Medicine 23*(4), 369–71.

Singer, J. and M. Small (1994). *Correlates of War Project: International and Civil War Data, 1816-1992*. Inter-University Consortium for Political and Social Research Ann Arbor, Michigan.

Tullock, G. (1971). The paradox of revolution. *Public Choice 11*(1), 89–99.

Ward, M. D., B. D. Greenhill, and K. Bakke (2010). The perils of policy by p-value: Predicting civil conflicts. *Journal of Peace Research 47*(4), 1–13.

Weidmann, N. B. (2009). Geography as motivation and opportunity: Group concentration and ethnic conflict. *Journal of Conflict Resolution 53*(4), 526–543.

Weidmann, N. B. (2011). Violence from above or from below? the role of ethnicity in bosnia's civil war. *Journal of Politics 73*(4), 1178–1190.

Weidmann, N. B., D. Kuse, and K. S. Gleditsch (2010). The geography of the international system:the cshapes dataset. *International Interactions 36*(1), 86–106.

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

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

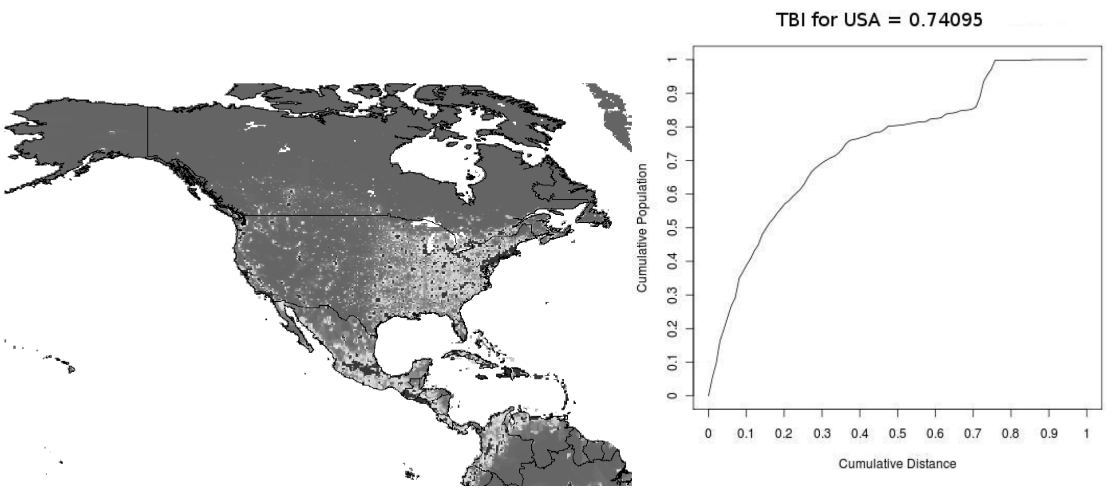

Figure 3: TBI curve for the United States. The curve on the right describes what fraction of the US population is within a certain distance from Washington D.C. The TBI expresses the area under the curve divided by the whole area of the plot. With the United States' population balance in the east, the resulting TBI is comparatively high. On the left, a visualization of the GPW dataset (CIESIN, 2005) for the region is visible.

Table 1: Regression results for outcomes and casualties based on Lyall and Wilson (2009) and the TBI.

	Dependent variable:								
	Outcome ordered logistic					Casualties negative binomial			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
REGIME	0.108**		0.110**	0.109**	0.105**	−0.022		−0.029	−0.046*
	(0.051)		(0.051)	(0.050)	(0.048)	(0.028)		(0.027)	(0.025)
SUPPORT	0.761**		0.806**	0.731*	0.806**	1.022***		1.101***	1.006***
	(0.388)		(0.386)	(0.385)	(0.372)	(0.217)		(0.220)	(0.196)
POWER	−0.630***		−0.637***	−0.415*	−0.769***	0.186		0.106	
	(0.229)		(0.229)	(0.219)	(0.265)	(0.200)		(0.196)	
ENERGY	−0.213		−0.207	−0.228		−0.120		−0.039	
	(0.214)		(0.214)	(0.215)		(0.110)		(0.108)	
OCCUPATION	2.291*		2.166	2.592**	1.950	−1.469**		−1.781***	−1.251**
	(1.339)		(1.332)	(1.280)	(1.186)	(0.654)		(0.663)	(0.580)
ELEVATION	−0.00004		−0.00004	0.00002		0.0001		−0.00004	
	(0.0004)		(0.0004)	(0.0004)		(0.0002)		(0.0002)	
COLDWAR	−1.278**		−1.246**	−1.269**	−1.124**	1.837***		1.671***	1.591***
	(0.571)		(0.570)	(0.563)	(0.557)	(0.333)		(0.328)	(0.323)
MECH.	−0.185		−0.146	−0.198		0.524***		0.482***	0.556***
	(0.276)		(0.274)	(0.276)		(0.165)		(0.160)	(0.148)
HELI.	1.947**		1.934**	1.886**	1.805**	0.490		0.365	
	(0.874)		(0.872)	(0.878)	(0.828)	(0.477)		(0.458)	
LANG.	0.065		0.061	0.075*	0.068*	−0.034		−0.021	
	(0.043)		(0.043)	(0.043)	(0.041)	(0.024)		(0.023)	
log(TRADE)	−0.584		−0.583	−0.590	−0.550*	−0.360*		−0.439**	−0.380**
	(0.406)		(0.407)	(0.402)	(0.329)	(0.216)		(0.209)	(0.157)
log(POPULATION)	−0.085		−0.104	−0.242		−0.333		−0.199	
	(0.154)		(0.155)	(0.155)		(0.205)		(0.202)	
DISTANCE	0.001*	0.001*	0.001**		0.001**	−0.0001		0.0002	
	(0.0004)	(0.0003)	(0.0004)		(0.0004)	(0.0004)		(0.0004)	
TBI		−1.134***	−1.466***	−0.984***	−1.890**		91.928***	81.573***	78.032***
		(0.136)	(0.045)	(0.048)	(0.855)		(26.789)	(21.627)	(20.899)
TBI^2							−65.922***	−60.447***	−56.850***
							(19.747)	(16.033)	(15.454)
(CONSTANT)						8.224***	−21.023**	−20.260***	−21.495***
						(2.422)	(8.978)	(7.266)	(7.030)
(CUTPOINT 1 2)	1.695***	−1.5343***	0.6891***	−0.9130***	1.9283				
	(0.1208)	(0.2002)	(0.1153)	(0.1682)	(1.0195)				
(CUTPOINT 2 3)	4.8559***	0.5407	3.8701***	2.1850***	5.0834***				
	(0.5365)	(0.2873)	(0.5411)	(0.5198)	(1.2480)				
Observations	65	65	65	65	65	57	57	57	57
In-sample CV	0.4462	0.5231	0.4615	0.4308	0.4308	855.2186	828.004	636.5732	697.6656
Out-of-sample CV	0.6	0.5231	0.6154	0.6	0.5692	1430.658	860.0593	1080.155	851.2102
θ						0.802***(0.130)	0.478***(0.074)	0.964***(0.159)	0.857***(0.139)
Akaike Inf. Crit.	131.38	141.52	133.13	132.96	126.12	1, 260.15	1, 281.79	1, 252.25	1, 247.12

Note: *p < 0.1; **p < 0.05; ***p < 0.01

Figure 4: Results from the empirical analysis of outcomes and severity of insurgencies. Five Ordered Logisitc and four Negative Binomial regression models were estimated to test the relevance of the proposed indicator.

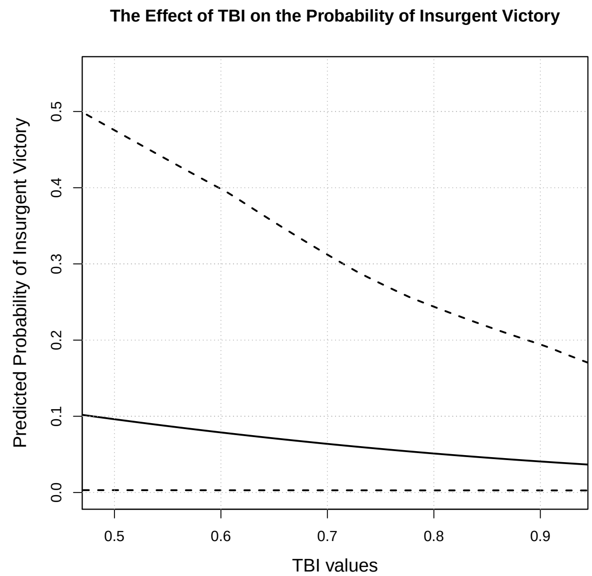

Figure 5: Predicted probabilities for the average country from model 4. The empirical model predicts a substantial decline for the probability of insurgent victory as a function of territorial balance.

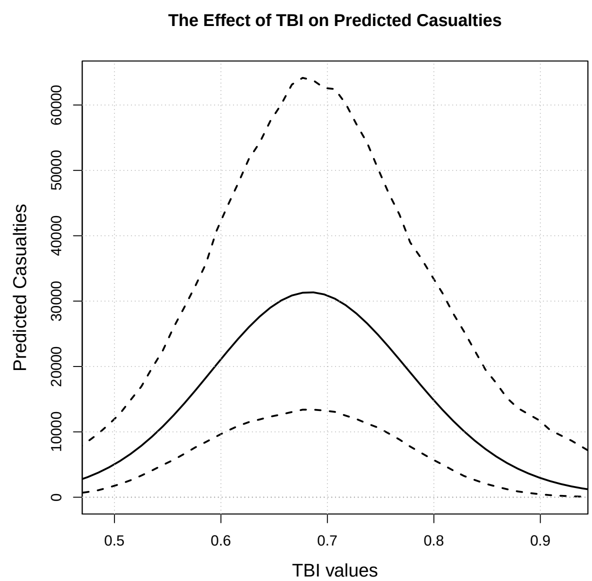

Figure 6: Predicted probabilities for the expected casualties as a function of TBI based on model 8. The dotted lines represent 95% confidence intervals.