Geographic Determinants of Indiscriminate Violence in Civil Wars

Sebastian Schutte[*,†,‡]

Zukunftskolleg and Department of Politics and Management, University of Konstanz

Abstract

What determines the type of violence used by military actors in civil wars? Drawing on Kalyvas' (2006) "information problem" and Boulding's (1962) "Loss of Strength Gradient", this paper proposes a simple model of how the violence becomes more indiscriminate as a function of distance from the actors' power centers. The proposed mechanism is a growing inability of the actors to distinguish between collaborators of the adversary and innocent bystanders. Tested on the conflict event level for 11 cases of insurgency, the results indicate that a simple distance-decay mechanism can explain the occurrence of indiscriminate violence to a large extent.

Keywords: Civil War, Conflict Events, GIS, Indiscriminate Violence, Afghanistan,

1 Introduction

The study of microdynamics of violence in civil wars has gained strong momentum in the past decade. Several new and revamped explanations have been proposed for why military actors in civil wars apply violence selectively or indiscriminately. The former type of violence is targeted at combatants, officers, and informants of the adversary, and the latter affects civilian bystanders. While this distinction has been made before (see: Mason and Krane, 1989), Kalyvas' (2006) widely cited study brought increased scholarly attention to the study of violence in civil wars, resulting in numerous publications (for example: Downes, 2007, Lyall, 2009, Lyall and Wilson, 2009, Condra and Shapiro, 2012).

Two strands of literature have regularly employed the conceptual distinction of selective and indiscriminate violence: Studies that seek to explain under which conditions civilians are intentionally harmed have focused on varying levels of military control (Kalyvas, 2006), initial motivations of the actors Weinstein (2007), and competition over local resources Metelits (2010). Studies that seek to explain the effects of indiscriminate violence on the conflict process follow either hypotheses of alienation or deterrence resulting from the use of indiscriminate violence (Downes, 2007, Lyall, 2009, Lyall and Wilson, 2009, Linke et al., 2012). With an overwhelming focus on the conflict process leading to civilian victimization or civilian victimization affecting the conflict process, too little attention has been given to the exogenous determinants of the type of violence used in civil wars.

Drawing on a simple distance-decay model from the power projection literature (Boulding, 1962) and Kalyvas' (2006) theory of indiscriminate violence, this paper shows that geographically operationalized power centers for both insurgent and incumbent are reliable spatial predictors of indiscriminate violence. Borrowing from Boulding's terminology, I refer to this phenomenon as the "Loss of Accuracy Gradient" (LAG). To empirically test the theory, this study combines detailed insights into the war in Afghanistan with the wide coverage of a cross-conflict sample of violent events drawn from 10 additional cases of insurgency in African countries. The results clearly indicate that military actors apply more indiscriminate violence as the distance to their power center increases. The next section will review the existing literature in more detail. After that, theoretical expectations are derived and tested in a large-N analysis.

2 Existing literature

The scholarly understanding of the driving forces behind indiscriminate violence has improved considerably in the last decade. Spearheading the recent turn towards the analysis of the microdynamics of civil wars, Kalyvas (2006, 69;149) convincingly argues that the scarcity of information and individual vulnerability in zones under enemy control leads to the application of indiscriminate violence. His theory of selective violence, however, assumes a more complicated mechanism to be at work: Military actors use violence in zones of predominant, but incomplete control to enforce collaboration and deter against defection. This leads to the empirical expectation that levels of selective violence should be highest in moderately contested zones of control, but not in areas of complete control or highest contestation. Focusing on the internal structure of military actors, Humphreys and Weinstein (2006) found that organizations that had to rely on material incentives to motivate their combatants and that lacked the ability to punish indiscipline were more likely to apply indiscriminate violence. Along these lines, Weinstein (2007) showed that the behavior of military actors toward the civilian population is affected by their initial motivations: Ideologically motivated rebels are less likely to engage in acts of indiscriminate violence than materially motivated ones. Offering yet another perspective, Metelits (2010) analyzes how violence becomes more indiscriminate when military actors compete over control and resources within the same territory.

In an in-depth analysis of the microdynamics of the Vietnam War, Kalyvas and Kocher (2009) showed that locations of selective and indiscriminate violence tend to be separated spatially, lending additional support to their control hypothesis. Bhavnani et al. (2011) extended Kalyvas' (2006) theory into a three-actor setting and found that it accurately predicted episodes of violence from the first and second intifada in Israel.

A noteworthy aspect of these theories is their focus on factors endogenous to the changing levels of military capabilities, such as territorial control, effective policing, and the availability of information about civilian loyalties.[1] While this perspective has generated many important insights, it also puts researchers and practitioners in a rather weak position to predict or anticipate episodes of indiscriminate violence, as it generally neglects conditions exogenous to the conflict process. Putting exogenous geographic conditions in the foreground provides

---

[1]Kalyvas (2006, 132-140) does acknowledge the importance of geographic factors for explaining the spatial distribution of military control and provides a comprehensive overview of qualitative insights into this relationship. However, his stylized theoretical model assumes violence to be a direct function of military control and to also follow its variation over time (page 119). But this notion ignores that fact that military control over any conflict zone is likely in the hands of the closest military power center for major parts of the conflict.

more empirical leverage for both inference and prediction.

This resulting perspective is well in-line with insights from the literature on revolutionary warfare and counterinsurgency theory. 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. Galula (1964, 23) discussed geographic constellations that would aid or hinder insurgent activity in wartime and particularly stresses the importance of remote base areas for insurgent movements. Similarly, (McColl, 1969, 5) pointed to inhospitable terrain that would allow for the build-up of rebel bases that could later be linked to form an "insurgent state". In his recollection of the Algerian war of independence, Aussaresses (2002, 26) summarizes that "[the] FLN could easily intimidate the countryside, but was having difficulty organizing itself within urban areas. That was precisely were terrorism was going to develop". Scott (2009) notes that the role of remote regions in providing safe havens for insurgents is usually well understood on a case-by-case basis, but is rarely communicated as a wider pattern:

> The [9/11] commission was quite aware that it was the combination of geographical remoteness, forbidding terrain, and above all, the relative absence of state power that made such areas recalcitrant to the exercise of power by the United States and its allies. What they failed to note was that much of the existing population in such areas of sanctuary were there precisely because these areas had historically been an area of refuge from state power (Scott, 2009, 127).

Center-periphery relationships and the conflictual potential inherent to them has inspired macro-historical explanations for war onset (Herbst, 2000), but the most basic relationship between the spatial constellation of military actors and the type of violence they use against one another was proposed by Arendt in her reflections on violence: "Those who oppose violence with mere power will soon find that they are confronted not by men but by men's artifacts, whose inhumanity and destructive effectiveness increase in proportion to the distance separating the opponents" (Arendt, 1970, 53).

Arendt's (1970) conjecture is reminiscent of a widely cited distance-decay model from the classic literature on power projection; Boulding (1962) assumed that a state's ability to project power was dependent on both its military strength and the distance to the theater of war. His notion of a "Loss of Strength Gradient" (LSG) predicts that for every unit of distance, a certain number of personnel had to be subtracted from the fighting forces and added to the supply troops. In theory, putting this relationship into numbers allows for the exact calculation of the limits in power projection. While heavily employed in the context of interstate conflict re-

search (for example Lemke, 1995), the model has also been modified to serve in the context of civil war research (Herbst, 2000, Gleditsch and Ward, 2001, Buhaug and Gates, 2002, Herbst, 2004, Buhaug and Gleditsch, 2006, Gleditsch, 2007, Cederman, 2008, Buhaug et al., 2008, Buhaug, 2010).[2] In the next section, I will combine Arendt's insights with Boulding's model under the notion of a "Loss of Accuracy Gradient".

3   The Loss of Accuracy Gradient

By assuming military control to be endogenous to the conflict process, contemporary theories deprive themselves of the opportunity to predict zones of increased indiscriminate violence ahead of time. Filling this gap, I demonstrate that political violence applied over large distances or into unknown geographies tends to become less selective: While highly selective close to the actors' power centers, it becomes less discriminate as it is projected across space. If this basic effect is understood, a sole focus on military actors' strategy and decision making seems insufficient to explain the internal dynamics and aggregate properties of civil wars. But how is it possible that spatial constraints so crucially affect the behavior of political actors? A discussion of the technological, tactical, and cognitive mechanisms involved follows.

3.1   Technological aspects

From a technological point of view, incumbent violence follows a clear trade-off between range and accuracy. An intercontinental nuclear missile cannot 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. While highly indiscriminate, aerial bombardment can be applied globally. Individual arrests and targeted assassinations are prototypical examples of selective violence, and their application is mainly confined to the areas under the actors' control. These technological constraints are somewhat symmetrical for incumbent and insurgent. For an insurgent movement, the available forms of violence also range from selective to indiscriminate – from targeted assassinations of individuals, to hit-and-run attacks on strategic targets with the possibility of civilian casualties, to globally applicable terrorism targeting mostly civilians.

---

[2]Buhaug (2010) offers compelling empirical insights that suggest that the LSG model can be applied to the study of civil conflict and summarizes the contemporary empirical literature.

The advancement of arms technology has only superficially contributed to solving this trade-off. Today, manned and unmanned aerial vehicles are 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. In measurement, accuracy refers to a system's ability to reflect a true value, while precision refers to the system's capability to reproducibly yield values with a constant offset to the true value. While it is hard to disentangle precision and accuracy in conflict event data, the conceptual difference can be highlighted with the dart game analogy: Throwing darts accurately means to 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 small distances between the darts. With regard to the application of violence in irregular conflicts this distinction proves decisive: An accurate application of force would reliably affect actual enemy combatants and would be conditional on the severity of their actions 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. Clearly, technological advancements along the lines of guided munitions have enabled violence to be applied with greater precision, but not nearly as much with greater accuracy. With regard to air strikes carried out by modern unmanned aircraft, high precision due to computerized targeting can certainly be assumed. At the same time, low accuracy is apparent in the fact that the vast majority of these attacks are so-called "signature strikes", i.e. directed against individuals who are presumed to be affiliated with insurgents, but whose identity is not known (see Heller, 2013). This example also illustrates that the technological mechanism underlying the LAG is not restricted to interstate conflict, but applies to civil war as well. As Kalyvas and Balcells (2010) observe, so called "symmetric non-conventional" and "conventional" tactics play an increased role after the end of the cold war in civil conflicts. In their empirical analysis, Kalyvas and Balcells (2010) found that access to heavy weaponry for rebel forces early on in civil conflicts had significantly increased after the end of the cold war.

3.2 Tactical incentives

Even if actors decide to not use aerial attacks or terrorist measures over long distances, they are presented with a tactical trade-off that generates incentives for indiscriminate violence outside their areas of control. Gener-

ally, combatants that advance into unknown and enemy-controlled territories will find themselves exposed to an increased risk of ambush and attack, leading to high casualties among the advancing troops. One way of mitigating this tactical disadvantage is the excessive use of firepower against *suspected* enemy positions. According to the motto of "spending shells, not men", this approach became a standard procedure in the Vietnam War (Greiner, 2009, 38). Especially when terrain conditions made patrols on foot hazardous and time consuming, "harassment and interdiction fire" replaced close quarter engagement with enemy combatants (Greiner, 2009, 150; Sheehan, 1988, 108; Lyall, 2009). Similarly, incoming sniper fire from civilian villages often lead to troops calling in "close air support", i.e. an airstrike against the entire village (Sheehan, 1988, 107). Clearly, random shellings or airstrikes in response to small arms fire qualify as indiscriminate violence.[3] These measures are not applied randomly in all conflict zones, but in those areas where inhospitable terrain makes rebel presence more likely and government presence is insufficient to tell apart rebels and bystanders. In South Vietnam, specially designated "free fire zones" were established in areas of insurgent control towards the northern border. Indiscriminate shellings in remote areas of a country and free fire zones might be considered a way of waging war against insurgents with minimal risk to incumbent forces. In areas of predominant state control, such as major cities, the pursuit of individual insurgents and their supporters is much less costly in terms of incumbent casualties. Moreover, the use of firepower is also much more restricted in comparison to remote regions. While these considerations apply to regular armies, there is a symmetrical effect for irregular forces with regard to deteriorating tactics as a function of distance. In remote, rural environments where insurgents enjoy superior control, identifying and punishing state collaborators can be performed on an individual basis. Outside these areas, insurgents can usually move and prepare attacks, but the network of civilian informants that they can otherwise rely on becomes thinner and less reliable. Direct attacks on incumbent forces are also extremely hazardous as retreating into difficult terrain is troublesome. As a result, insurgents predominantely attack the centers of state power by indiscriminate means. Terrorist tactics that allow for time-delayed explosions are a classic example of long-distance insurgent violence. Again, there is the same trade-off between distance and accuracy.

---

[3]Lyall (2009) reports that this tactic was also used more recently in the Chechen wars.

3.3 Cognitive constraints

Above and beyond the technological constraints and tactical incentives, there is a cognitive dimension that makes indiscriminate violence more likely in areas unknown to the actors. The problem of limited information about civilian loyalties and the resulting incentives for indiscriminate violence have been discussed before, both from a psychological as well as rationalist point of view (Kalyvas, 2006, 69). The psychological mechanism at work, according to Kalyvas, is a fundamental distrust towards the civilian population, as well as frustration, uncertainty, fear, and anxiety resulting from the inability to tell apart foe and bystander (see also: Greiner, 2009, 124). Moreover, insurgent tactics of surprise attacks and timely retreats make the clear identification of attackers extremely difficult. As a result, the local population is suspected of being actively involved with the uprising. This impression of meeting civilian resistance in pointless pursuit of an invisible enemy is commonplace in counterinsurgency campaigns: Ellsberg summarized the daily routine of US troops in the Mekong Delta as follows: "Foreign troops far from home, wearing helmets and uniforms and carrying heavy equipment, walking along dikes in formation and getting shot at every half hour mostly by ragged local irregulars firing from tree lines that bordered their homes" (Ellsberg, 2003, 167). Similarly, Sheehan described the problem: "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). Greiner (2009, 35) rightfully points out that the lines between objective threat and subjective impression might have been blurry at times, but both of them contributed to heavy handed actions.

Especially in out-of-area operations, fast rotation cycles add to the inability of combatants and commanders to learn about local loyalties and heuristics for discrimination, as discussed by Cowper-Coles (2011, 167) for the contemporary war in Afghanistan. By transferring commanders out of the theater of operation on a six-month basis, the ability to learn vital lessons about civilian loyalties is greatly reduced (Cowper-Coles, 2011, 62). Clearly, such deployment schema greatly add to the fundamental problem of fighting population-centric campaigns when knowledge of the population is missing. But even the goal of becoming an expert on local ties as suggested by the US Army Counterinsurgency Field Manual (DOD, 2007, Appendix A, Paragraph A2) seems ambitious at best for actors operating outside their cultural circle of origin. Many historians and sociologists

have struggled to paint a coherent picture of the Afghan society and their most successful attempts require the introduction of social units and concepts that are foreign to most Westerners (Dorronsoro, 2005). Mental access to foreign regions is hard to acquire and difficult to communicate. This problem is somewhat amplified by the particularities of irregular war, in which asymmetry in numbers of combatants is a defining feature (see: Fall, 1965, 139-140). Put differently, in asymmetric conflicts it is not only hard to tell combatants and civilians apart, but there is also only a small number of fully committed irregular combatants to be identified, making the pursuit all the more frustrating (Crawford, 1958, 56). Combining this with the fact that a false negative identification of rebels puts combatants in mortal danger generates incentives to "shoot first and ask questions later".

The scarcity in information that the military actors have to rely on in identifying enemy collaborators provides an ideal breeding ground for simplifying heuristics. Logically, knowing that insurgents are present in a certain location does not imply that everybody present at that location is an insurgent. But the association of places to loyalties fills the void left by what Kalyvas (2006, 148) calls the "information problem". The term "Indian Country" illustrates the general suspicion of US troops towards the entire civilian population in the northern provinces of South Vietnam (Greiner, 2009, 145). Together with "Enemy Central", the term has recently been revamped to describe Kunar province in Afghanistan and Babil province in Iraq. Cultural and linguistic barriers in out-of-area deployments are the most obvious cognitive constraints to understanding local loyalties, but the effect of scarce information leading to excessive brutality based on simple heuristics is considerable even in purely domestic conflicts. Again, the accessibility of information, situation awareness, and the suitability of heuristics for identifying civilian loyalties are functions of distance. Close to the military power centers, combatants assume that civilians benefit from the protection they can provide. As distance to the power centers grows, so does the uncertainty about the presence of enemy combatants and collaborators. Close to the enemy power center, the presence of enemy combatants and informers can be assumed. As simple cognitive heuristics offer solutions to the uncertainty of who is fighting for whom, indiscriminate violence in remote regions becomes more likely in terms of trigger-happy reactions to suspected enemy presence.

3.4 A distance-decay model

The conclusions from this theoretical discussion are straightforward: Violence declines in selectiveness as a function of distance from the actors' power centers. Modifying Boulding's (1962) "Loss of Strength Gradient"

into a "Loss of Accuracy Gradient", I therefore assume the *quality* of violence to deteriorate as a function of distance instead of its *quantity*. Within this model, the initial levels of accuracy or discriminatory potential are not necessarily the same for all actors and the gradient in the declining quality might also vary. For example, modern means of surveillance or a better intelligence apparatus might very well lead to an increased ability of the actors to identify enemy combatants and collaborators, but this ability to apply violence accurately is expected to decline with growing distance to their power centers.

Operationalizing power centers as spatial locations is necessary to test this hypothesis. For the state, the capital city naturally qualifies as its center. In most irregular civil wars, capturing the capital city is also the final objective of the uprising.[4] For the rebels, the situation is less obvious. McColl (1969) describes several determinants for the emergence of rebel bases in irregular conflict, such as distance to international borders and terrain accessibility. While controlling for these factors, I assume the rebel power center to generally emerge in areas most remote from the capital city. Using international borders for retreat and supply, insurgent movements tend to use remote areas to build up of their bases (Salehyan, 2009). Therefore, I operationalize the insurgent power center as the area of a war-torn country that is most remote from the capital city while still being affected by conflict. Clearly, such assumptions about the locations of rebel strongholds cannot be made for all civil wars. For example, the Yugoslav civil war mainly split the country along ethnic lines and not in terms of a territorial divide that arose from a popular rebel movement challenging the state. In such cases, the pre-war settlement locations are much better predictors of violence than distance to the former capital city (Weidmann, 2011). Therefore, both the theoretical scope of the argument as well as the case selection are restricted to irregular civil wars. The theoretical expectations naturally lead to the following hypotheses:

H1: Distance to the capital city has a positive effect on the probability of incumbent violence being indiscriminate.

H2: Distance to the capital city has a negative effect on the probability of insurgent violence being indiscriminate.

H3: Areas closest to and farthest away from the capital city see the highest levels of indiscriminate violence.

---

[4]In the empirical analyses, I also control for distances to major cities which can be assumed to be under incumbent control for most of the conflict. Unfortunately, data on the locations of military bases (i.e. the most appropriate proxy for the centers of state power during civil conflicts) are not publicly available.

While the scope of the argument is limited to insurgencies, it provides an added value by focusing on exogenous conditions that affect the spatial distribution of indiscriminate violence in civil wars. In doing so, the approach does not fundamentally challenge the insights generated by others, but it allows for central components of their theories, such as Kalyvas' "information problem", to be associated with geographic conditions. The advantage of this approach is that it only requires very modest and general assumptions to predict an association between the spatial configuration of a theater of war and the predominant types of violence that occur. The disadvantage is that the measurement is compatible with several causal stories, all of which begin with the actors extending their operations beyond their power centers and end with them resorting to indiscriminate measures. The next section discusses how the theory will be tested.

4 Empirical strategy and case selection

The increasing availability of conflict event data has generated unprecedented possibilities for analyzing the subnational characteristics of civil wars. Such data have been used in a large and rapidly growing number of publications on the microdynamics of civil war (Hegre et al., 2009, Schutte and Weidmann, 2011, Weidmann and Salehyan, 2012, Braithwaite and Johnson, 2012). While most of these studies focus on single conflicts, a series of studies have also attempted to find generalizable patterns in event data across conflicts (Buhaug and Rød, 2006, Bohorquez et al., 2009, Buhaug, 2010). However, a decisive factor sets these studies apart from the empirical analysis of the introduced LAG: Instead of modeling the presence or absence of violence, the presented theory focuses on the type of applied violence. This distinction calls for an empirical record that reveals the type of violence as well as its location. Moreover, acts of violence must be attributable to one of the military actors in order to test the proposed distance-decay mechanism. Unfortunately, such data are hard to come by.

The Armed Conflict Location and Event Dataset (ACLED) (Raleigh and Hegre, 2005) provides geocoded information on acts of violence, but it does not differentiate between civilian and military casualties. While certainly useful in other cases, these data do not lend themselves to the analysis of variations in the quality of violence as a function of location. More recently, a conflict event dataset covering African civil wars between 1990 and 2010 has been released that provides information on both civilian and military casualties (Sundberg

et al., 2011). The Georeferenced Event Dataset (GED) exclusively codes lethal encounters in civil wars. Counts for civilian and military fatalities are given for each conflict event. A caveat for the problem at hand is that conflict events cannot be attributed to one of the military actors. This is understandable from a data collection point of view as the coding relies on media sources. For journalists, competing claims about the initiation of violence are often hard to verify and the affiliation of the perpetrators is usually the most contested part of any story emerging from the turmoil of civil war. Consequently, GED does not code the initiator of violence. Beyond the realm of basic research, data on conflict events have also been collected by armed forces engaged in counterinsurgencies. The most recent data collection system for the microdynamics of ongoing conflicts within the US military is called SIGACT, an abbreviation for "significant activity". SIGACT files are referenced in time and space, indicate the specific type of incident, and record casualties and the initiator of the event. These data have been used in a responsible manner for basic research in other recent publications (O'Loughlin et al., 2010, Linke et al., 2012, Condra and Shapiro, 2012, Zammit-Mangion et al., 2012, Braithwaite and Johnson, 2012, Carpenter et al., 2013, Weidmann, 2014, Schutte and Donnay, 2014). The SIGACT data provide all necessary information for testing the proposed hypothesis. It is nevertheless restricted to two cases – Afghanistan and Iraq – and only Afghanistan clearly qualifies as an insurgency, while the war in Iraq blends elements of an insurgency with ethnic and communal violence (Weidmann and Salehyan, 2012).

Given these restrictions in data availability, I decided to combine the detailed insights that can be gained from SIGACT with the larger coverage of GED in two separate empirical analyses. In a first step, I conduct an analysis of the SIGACT data from Afghanistan. In this case, the initiator of violence is clearly coded. However, the casualty figures from SIGACT are not trustworthy, as soldiers in the field have strong incentives to downplay the levels of civilian casualties resulting from their actions. Moreover, confirming casualties is easier in certain scenarios than in others. Calling in an airstrike against a suspected enemy position might very well result in civilian casualties that go unnoticed, for example. As a result, a related study on civilian casualties in Iraq relied on external casualty counts (Condra and Shapiro, 2012). Another problem with SIGACT is that it is not a complete data collection of all conflict events: the activities of non-US ISAF troops or other arms branches, such as the US Air Force are only sporadically recorded. To circumvent these problems, I operationalize selective and indiscriminate violence within the SIGACT study not in terms of the alleged casualties, but in terms of the type of violence employed by the actors. I also only focus on events that are frequently reported.

After that, I test for the generalizability of the theory in an analysis of the spatial distributions of civilian and military casualties. Since the GED dataset does not code the initiator of conflict events, instances of indiscriminate violence cannot be clearly attributed to actors, but the locations that see higher levels of indiscriminate violence can be analyzed across cases. In summary, the empirical strategy seeks to combine the detailed insights of SIGACT with the wide coverage of GED.

4.1   The Loss of Accuracy in Afghanistan

The ongoing war in Afghanistan was identified as a case of insurgency that lends itself to testing hypotheses 1 and 2, i.e. the distance decay in the quality of violence for both military actors. What makes Afghanistan a suitable case? Its general socio-economic conditions, the structure of the uprising, the international context, and the sequence of events at the macro-level are all typical of this type of conflict (for selecting typical cases to craft general insights see Seawright and Gerring, 2008, 297). Moreover, the SIGACT dataset covering the war in Afghanistan has one distinct advantage over other data sources: it codes the initiator of violent attacks explicitly.

From a socio-economic point of view, the country is a risk candidate for civil war. Widespread poverty, a weak central government, a recent history of intense political violence, forbidding mountainous terrain, and a patchwork of intermingled ethnic groups with varying access to political power and wealth make Afghanistan a prime candidate for civil war with regard to the central variables associated with war onset (Collier, 2000, Buhaug and Gates, 2002, Fearon and Laitin, 2003, Cederman et al., 2010, 2011). A downside for testing the proposed theory, however, is the fact that SIGACT data on Afghanistan covers conflict events that involved US troops that were deployed outside their home country. Consequently, language barriers and cultural differences between incumbent forces and the local population in Afghanistan might have been greater than in other insurgencies. This would make Afghanistan an eligible, but not perfectly typical case to study.

The structure of the uprising is neverthelss representative for conflicts of this kind, combining an irregular local insurgency with the build-up of a shadow administration of judges and tax collectors. Moreover, clandestine international support for the insurgency is assumed to take place and can be found across a variety of cases, such as Chechnya and Ingoshetia (Moore and Tumelty, 2008), Vietnam (Sheehan, 1988, 650), and, of course, the anti-Soviet insurgency in Afghanistan (Wright, 2007, 120). Apart from that, the sequence of macro-events

that led to a large-scale insurgency is also typical of a wider class of cases: A government is replaced through outside intervention and subsequent occupation of the country. The new government faces a problem of legitimacy and is heavily reliant on outside support. Elements loyal to the former administration start a protracted campaign to topple the new incumbent.

4.2 Explanatory variables

For the samples of observations from SIGACT and GED, I coded explanatory variables through a GIS procedure. Distances to the power centers were established based on the CShapes dataset (Weidmann et al., 2010). To account for other factors beyond this main independent variable, I coded the approximate terrain *elevation* from a global elevation model Gesch et al. (1999), *population* counts from CIESIN (2005), as well as the number of *previous conflict* events within a 50 kilometer radius less than one month prior to the incident. Furthermore, measures of wealth from a spatially referenced dataset on the subnational distribution of goods and services was included (*GECON*) (Nordhaus et al., 2006). Moreover, I added an estimate for the natural *landcover* (Hansen et al., 2000) at the site of the conflict event, as well as data on the effective traveling distance to the nearest city that had at least 50,000 inhabitants in the year 2000 (Nelson, 2008), hereafter referred to as *urban distance*.[5] Finally, I constructed a *line-of-sight* measurement to account for a tactical particularity: In areas with limited lines-of-sight due to natural obstacles, actors might rely on indirect fire instead of direct fire.[6] Since I coded the first type of attack as selective and the second type as indiscriminate, I needed to control for this factor. Using the digital elevation model by Gesch et al. (1999), I calculated the number of surrounding cells that are visible from any location. This calculation involved the specification of "horizon" in terms of a maximal distance from the cell under investigation (50 kilometers in this case). For all cells within that horizon, Bresenham's (1965) algorithm was used to calculate all cells along a straight line connecting the origin and the target cell. In a second step, elevation levels along this line were used to calculate angles between the cell under investigation and the cells along the line. The number of visible cells was then established by counting the number of cells along this line for which no steeper angle had been calculated for any preceding cell. In this way, cells along this line

---

[5]The landcover data was downloaded at the 1 km² resolution. and distinguishes among 13 landcover types ranging from urban to densely forested. Omitting water as a surface type, I use this variable as an ordinal measure for vegetation density. The urban distance data was downloaded at the 5 km² resolution. It combines georeferenced information on roads, railroads, navigable rivers, and land cover to compute approximate travel times to the nearest major city.

[6]Please refer to the supplementary information for a detailed explanation on how this measure was calculated.

with no obstructing cell in front of them were established. Since this procedure was repeated for all cells within the horizon, a count of all visible cells was established for each cell in the elevation dataset. The explanatory variables were available as georeferenced data with varying resolutions. In order to associate the explanatory variables with the conflict observations, I mapped the conflict events to their nearest geographic neighbors in the explanatory datasets. As a result, samples from GED and SIGACT could be analyzed in a multivariate regression analysis. Descriptive measures for all explanatory variables can be found in the appendix.

4.3   SIGACT Data

The version of SIGACT used here covers the time period from 2004 to 2010 and amounts to 76,247 records. All records are time- and georeferenced, coded as insurgent or incumbent activity, and distinguish between 154 types of events. SIGACT reports are passed up the chain of command, sometimes from the platoon level, allowing for an extremely detailed view of the conflict from the U.S. perspective. From a quantitative standpoint, the SIGACT data provide the most complete view on the Afghan War. As mentioned before, one shortcoming of this data is that they were coded by soldiers in the field that perceive the conflict subjectively and report violence events with strong incentives to attribute civilian casualties to the enemy, or to not report them at all. Therefore, I decided to operationalize indiscriminate violence as the use of heavy arms, as described in detail below. A cross-sectional view of the data reveals the existence of two major conflict zones as illustrated in figure 1. One zone centers around the city of Kandahar in the Helmand province while the other lies along the mountainous tribal areas bordering Pakistan. The northern and eastern parts of the country seem less affected by civil war, although single instances of violence are recorded in SIGACT.

4.4   Operationalizing indiscriminate violence in SIGACT

The operationalization of selective and indiscriminate violence in this section follows the premise of section 3. One can safely assume that several types of incidents in SIGACT are less selective than others. I focused on conflict events that are especially frequent in the empirical record to allow for the largest possible external validity of the analysis and to prevent the insights from being driven by marginal conflict episodes. Moreover, the selection of relevant events focused on types of events that are comparable for incumbent and insurgent.
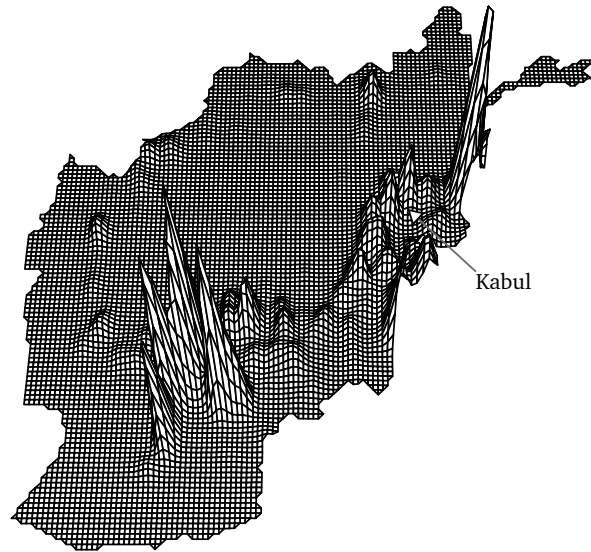
15

Figure 1: The figure above shows a (Gaussian) density estimate for the locations of conflict events in Afghanistan between 2004 and 2010. It gives an impression of the relative intensity of conflict in different parts of the country. Violence was most intense in Helmand province and in the tribal areas bordering Pakistan. The spike in the northeastern part of the country is the result of intense fighting in the Korangal valley.

Coded type of violence	SIGACT event category	N
Incumbent selective	direct fire	823
Incumbent indiscriminate	close air support, indirect fire	595
Insurgent selective	direct fire	15458
Insurgent indiscriminate	mine strike, indirect fire	7173

Table 1: Event categories for selective and indiscriminate violence in the SIGACT data

Table 1 gives an overview of the event categories associated with incumbent and insurgent violence. Please refer to the supplementary information for an overview of other categories in SIGACT and an additional discussion of my coding choices.

Indiscriminate insurgent violence

Mine strikes were counted as acts of indiscriminate violence. In order for an explosive device to classify as a landmine it must be victim-activated. It is this technical particularity that renders landmines extremely inaccurate, since no identification of the target is possible by the attacker. Similarly, indirect fire allows an attacker

16

to hit targets beyond his line of sight. Moreover, it allows the attacker to deliver heavy and explosive munitions over greater distances. These tactical characteristics come at a decisive cost in population-centric warfare: limited accuracy and high lethality. Effective indirect fire usually relies on an artillery spotter having line-of-sight contact with the target to report back to the shooter. It usually requires several iterations of shooting and re-aiming to hit a target. More importantly, explosive munitions destroy the homes and property of innocent bystanders, even if they do not physically harm civilians. These characteristics make *indirect fire* less discriminate than direct small arms fire and these events were therefore counted as indiscriminate.

Selective insurgent violence

Direct fire by insurgents has certainly claimed the lives of civilians, but it still provides a more selective way of targeting collaborators and incumbent forces than indirect fire. More importantly, due to line-of-sight contact being a precondition of the use of direct fire, insurgents at least know what or whom they are shooting at in combat situations.

Indiscriminate violence by US Forces

The rules of engagement of US Forces put restrictions on using lethal force. Although ISAF can probably use *indirect fire* more professionally than insurgent forces, the problem of high lethality combined with low accuracy remains. Even if measures are taken to spare the lives of bystanders, large-scale material destruction is still a natural byproduct of explosive munitions. Therefore, instances of *indirect fire* were counted as indiscriminate violence for the incumbent side. Generally, SIGACT does not contain information on Air Force activity in Afghanistan. This is due to the fact that soldiers on the ground file the reports and only sometimes include references to air strikes. If air strikes were carried out as part of other fighting activities, the incident might simply not be labeled as such. Nevertheless, *close air support* was counted as an instance of indiscriminate violence since it applies violence more destructively and less selectively than direct small arms fire.

Selective violence by US Forces

Direct fire was also counted as an instance of selective violence. Again, these events are frequent enough to allow for generalizable insights and arguments for counting direct fire of the insurgent side as selective; this

also applies to the incumbent side.

4.5   SIGACT results

To establish the effects of the distance to the power centers on the type of violence applied, I estimated logistic regression models for both insurgent and incumbent violence. The unit of analysis in these models is the conflict event. I reduced the sample of all SIGACT events to those that relate to violent incidents coded either as selective or indiscriminate according to table table 2 on page 36. Models 1, 2, and 3 predict indiscriminate violence for the insurgent side while models 4, 5, and 6 predict incumbent violence. Models 2 and 5 contain all explanatory variables, whereas 3 and 6 consist of subsets of the explanatory variables chosen to improve the goodness-of-fit as expressed in lower AIC values. AIC is a suitable statistic for measuring the benefit of adding more variables against the costs of adding random variability (Akaike, 1974). Models 1 and 4 do not contain the main independent variables (the distances to Kabul and the Pakistani border) and serve as a baseline for the AIC statistic. Comparing models 1 and 2 with regard to their AIC scores therefore allows us to see that the distance variables substantially improve the model fit instead of merely overfitting model 1. Similarly, a comparison of models 4 and 5 shows that the inclusion of the distance variables also lowers the AIC for the incumbent side, which further underlines the relevance of the variables.

In summary, the lower AIC values indicate that the distance variables substantially improve the model. They do not, however, show how the explanatory variables are associated with types of violence used by the actors. I will therefore discuss the results from a substantive point of view below. As first glance, table 2 shows a number of variables significantly associated with the types of violence used by the actors. These very strong results are certainly partially driven by the large number of observations – 22,512 events for the insurgent side and 1,288 events for the incumbent. But more importantly, the fact that these geographic variables are significant predictors of the types of violence indicates that locations of selective and indiscriminate violence tend to be separated in space (see also: Kocher et al., 2011), which is in line with the theoretical expectations.

Distance to Kabul, the main independent variable, is negatively associated with insurgent indiscriminate violence, but positively associated with incumbent indiscriminate violence. As argued above, a possible explanation for this effect might be that incumbent forces use more heavy weapons in order to secure and defend valuable areas and infrastructure to prevent them from falling into the hands of insurgents. An alternative ex-

planation is that much of Afghanistan's wealth is located in the border region to Pakistan, in terms of natural resources located there (Peters et al., 2007). Following hypotheses 1 and 2, this region should see higher levels of incumbent indiscriminate violence and lower levels of insurgent indiscriminate violence. Further analysis would be required to fully disentangle these two possible explanations.

Urban distance has a positive effect on insurgent indiscriminate violence and no significant effect for the incumbent side. At first glance, this result is contrary to the expectation that insurgents use indiscriminate violence in the cities. However, the specific coding of the variables associates indiscriminate violence with the use of heavy weapons, such as mortars. Therefore, the result is most likely driven by the fact that insurgents often refrain from using mortars and mines in the cities, restricting their use to more rural settings which offer better opportunities for retreat after an attack. This tactical particularity is less likely to drive the estimates on the incumbent side, as incumbent forces can rely on long-range artillery and air support and do not need to retreat after every engagement. A refined empirical analysis would be necessary to understand whether this effect is caused by the specific coding choices for the SIGACT categories or remains robust across proxies of indiscriminate violence.

The positive effect of natural *landcover* on insurgent indiscriminate violence can be interpreted along these lines as well: Insurgents lack the means to transport heavy weapons in the open. Given the efficiency of modern air surveillance, it might be necessary for insurgents to utilize landcover for transporting these weapons. Again, due to the clustering of insurgent activity along the Pakistani border, local conditions can drive the statistical inference. With many parts of the country sparsely forested, the Nangarhar province in eastern Afghanistan is an exception. A comparatively mild climate and a continuous fresh water supply from the Hindu Kush, the region is comparatively forested and engaged in timber production. However, it has also seen increased insurgent activity due to its proximity to the border. Again, additional testing would be requried to fully rule out this alternative explanation for the effect. The *previous violence* variable expresses the number of conflict events that took place up to one month prior to the event under investigation and within a 50 km radius. Interestingly, the estimate for the incumbent side is negative, while it is positive for the insurgent side. Asymmetry in the applied tactics is a candidate explanation for this effect: Insurgents tend to commit "hit and runs", i.e. leave the area after initiating an attack, while incumbent forces hold on to territory. As the battle progresses, the incumbent side increasingly relies on air and artillery strikes to defend its positions, while the insurgent side

relies more on mobile attacks. In summary, the variables *distance to Kabul, GECON, landcover,* and *previous violence* have opposite effects on incumbent and insurgent indiscriminate violence. The results for the main independent variable is in line with the theoretical expectations of hypotheses 1 and 2, but these results only reflect the conditions within one single conflict. In order to test the theoretical expectations more generally, I also conducted an analysis of irregular civil wars in Africa which will be discussed in the next section.

4.6   Civilian casualties in African insurgencies

To broaden the analysis beyond Afghanistan, I have analyzed a subset of the conflict events from GED. An overview of these cases is given in table 3. The GED analysis was restricted to cases of irregular war and identified through a dataset of insurgencies by Lyall and Wilson (2009). As mentioned previously, the wider coverage of the GED data comes at a price: Initiators of violence are not coded, and the type of violence for the conflict events is not disaggregated into tactical categories. As a result, the dependent variable of this analysis is not the type of battle event, but the number of civilian casualties that resulted from the application of violence.

Bearing in mind the theoretical discussion, the empirical expectations of this section are straightforward: The two actors employ violence indiscriminately close to the other actor's power center (H3). In terms of casualties, this would imply that civilian casualties are most likely to arise in two locations: close to the capital city where insurgent forces are more likely to engage in indiscriminate actions, and in the most remote regions where incumbent forces employ violence indiscriminately. This theoretical expectation translates into a U-shaped effect for the number of civilian casualties as a function of distance between the capital city and the nearest international border. The table below shows the cases of irregular civil war that are covered in GED:
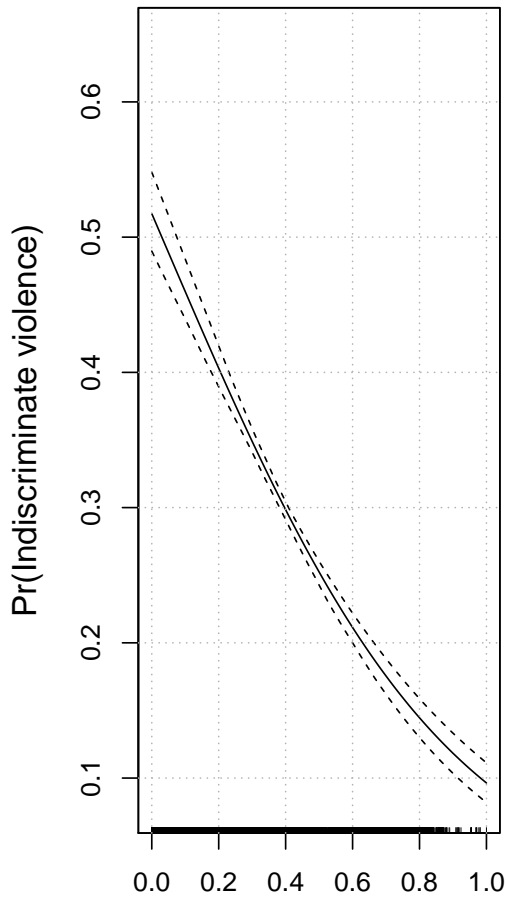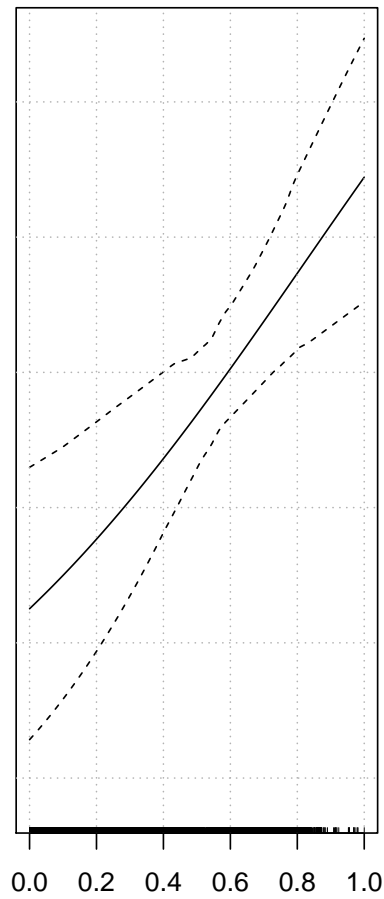
Figure 2: Plots showing the effect of the main independent variable on the probability of indiscriminate violence for both incumbent and insurgent. To generate these figures, *distance to Kabul* was systematically varied while all other variables were held constant at their means.

No.	Country	Scenario description	Period(s)
1	Algeria	Large-scale civil war after 1991 military coup	1991-
2	Burundi	Ethnic civil war of the Hutu against the Tutsi-dominated government	1994-2005
3	Chad	Southern armed groups challenged the political and economic dominance of the north	1994-1998
4	DR Congo	Uprising against President Mobutu followed by transnational proxy wars	1994; 1996-1998; 1994-1999
5	Djibouti	Ethnic insurgency of the Afar to achieve political participation	1991-1994
6	Guinea-Bissau	Uprising against President Vieira to enforce a change in government	1998-1999
7	Ivory Coast	Irregular war after mutinous soldiers picked up arms against the government	2002-2004
8	Liberia	Irregular civil war aimed at removing President Taylor from power	2000-2003; 2001
9	Rwanda	Ethnic civil war after the Hutu-dominated government was removed from power in 1994	1994-2002
10	Sierra Leone	Irregular civil war against the state followed by infighting amongst rebel organizations	1991-1999

Table 3: Overview of the GED cases under investigation based on the list of insurgencies by Lyall and Wilson (2009).

4.7 GED data

GED is a detailed data collection coded and maintained by Uppsala university. Covering lethal events from both civil wars and communal unrest, GED provides 20,396 observations of violence that took place between 1990 and 2008 in Africa. The GED data are based on an elaborate coding procedure that ensures reliability through redundant codings (Sundberg et al., 2011). Most importantly for this study, the GED dataset provides geographic coordinates along with all observations, allowing the hypotheses derived above to be tested empirically.

Since the theoretical scope of the argument is restricted to insurgencies where rebels usually operate from the periphery, a subset of conflicts in GED had to be identified that meet this criterion. I therefore used a dataset by Lyall and Wilson (2009) that is restricted to clear-cut insurgencies. Unfortunately, this dataset and GED only overlap partially, since the GED codes conflict events on the African continent that took place between

1990 and 2008, while the Lyall and Wilson (2009) dataset codes military outcomes of insurgencies on a global level from 1800 to 2010. As a result of this partial overlap, conflicts from only 10 countries appear in both data collections.[7] The resulting sample consists of 3,744 observations from the GED. This sample was generated by preserving only those GED observations that took place during and within the same country as the insurgencies coded by Lyall and Wilson (2009).[8]

Moreover, instances of communal violence and one-sided violence were excluded from the sample since their occurrence is beyond the scope of the theory.[9] Distances to the capital city were normalized for each country, using the most remote conflict event as the maximal possible distance. All observations were pooled and analyzed in the aggregate.

4.8   Operationalizing indiscriminate violence in GED

With regard to the GED dataset, the type of violence is more easily measured than in SIGACT: Since every observation reflects counts for military and civilian casualties, the numbers of civilian casualties directly reflect higher levels of indiscriminate violence. Therefore, the dependent variable in the GED study is the number of civilian casualties in violent events, while the dependent variable in the SIGACT study is the type of applied violence according to table 1.

4.9   GED results

In order to establish descriptive insights, I plotted the number of civilian casualties as a function of normalized distances to the capital city. As visible in figure 3, higher levels of violence against civilians can be observed for small and large distances, but the sample is dominated by a number of severe incidents at medium distances. The U-shaped relationship suggested by the theoretical discussion is more clearly visible, once extreme outliers are omitted (middle row in figure 3).

---

[7]In a robustness check, I also ran the analysis on the unrestricted sample including all GED events and found that the substantive effect also holds for the entire sample (see table 4). This test was necessary to show that I am not selecting cases that match the theory while omitting others.

[8]One unintended consequence of this case identification scheme is that the identified insurgencies in GED are all within Africa. The African state system clearly presents a special case within the global sample, as the corresponding borders were largely drawn during the 1884-85 Berlin conference. The generated insights might therefore not fully translate into cases where state formation took place under non-colonial conditions.

[9]Clearly, military actors that deliberately target civilians could also drive the casualty figures in GED and these types of attacks lie outside the presented theory and empirical analysis.
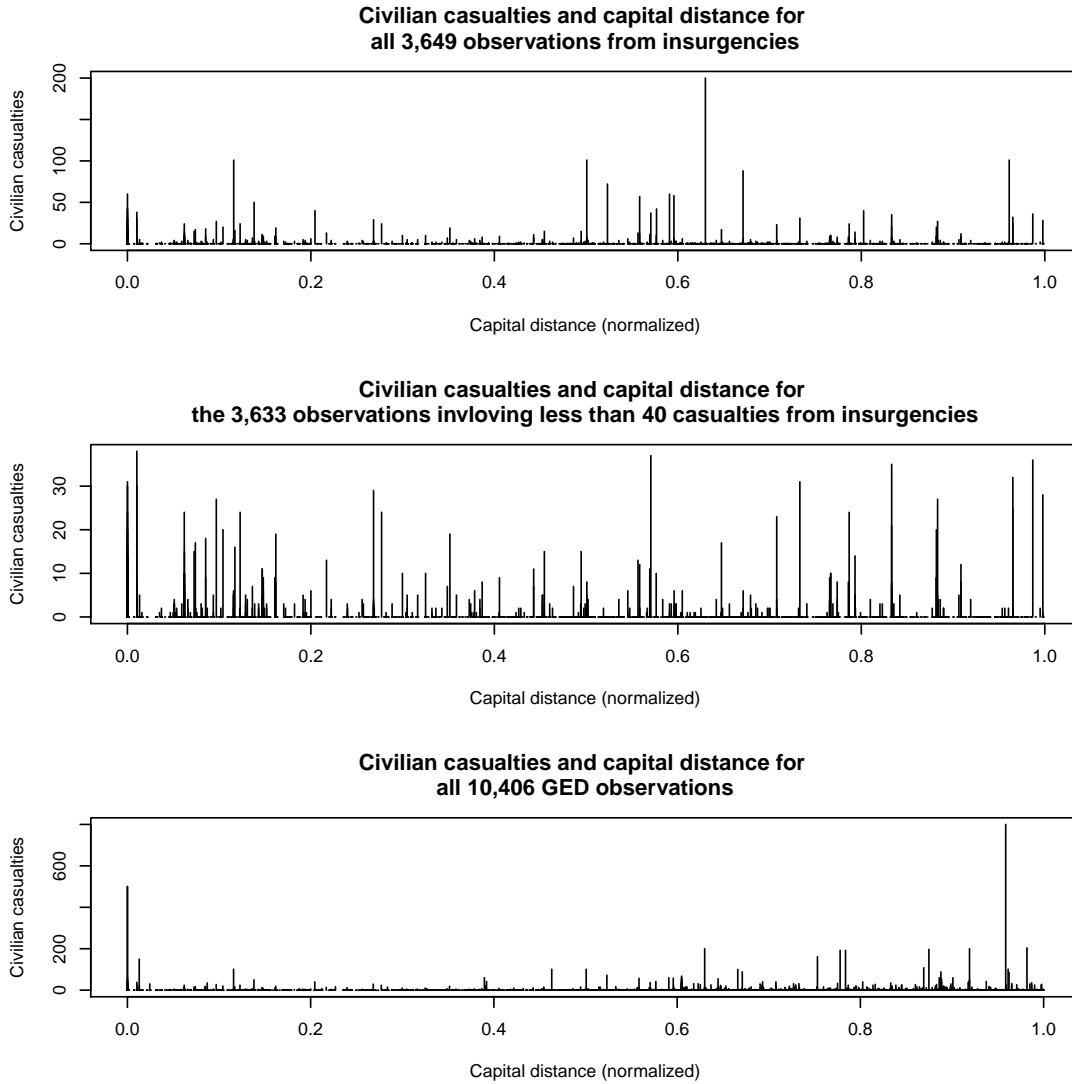
Figure 3: Casualties in the GED dataset as a function of distance from capital for all observations from insurgencies (above), less severe incidents from the insurgencies (middle), and all GED observations (below). Note the U-shaped relationship that is visible in the lower two figures.

In order to assess the external validity of the proposed mechanism, conflict events from 10 cases of insurgency in the GED (Sundberg et al., 2011) were analyzed with multivariate regression models. Corresponding results can be found in table table 4 on page 37. Although they do not reveal the initiator of violence, the GED observations code civilian and military casualties. Predicting the number of civilian casualties as a function of location allows us to test the proposed theory from another angle and with a considerably wider scope.

Model 7 presents only the main independent variable, and model 8 only the additional controls. Model 9 contains all explanatory variables and model 10 only those that contribute to a lower overall AIC score. Model 11 was run with the full sample of all civil war-related GED events, excluding only interstate war and one-sided violence. This model was included as a robustness check to rule out the possibility that the central insights were driven by the case selection. Model 12 contains a smooth term for the main independent variable. A first indication of the relevance of the normalized distance variable is the fact that the full model 9 features a lower AIC score than model 8, which contains all but the main independent variable. Since indiscriminate violence cannot be attributed to one specific actor in the GED, only the expectation that the capital city and the periphery should see the highest levels of indiscriminate violence can be tested. Hypothesis 3 expresses the corresponding empirical expectation of a quadratic (U-shaped) effect for the normalized distance from the capital city. As visible in table 4 on page 37, the quadratic specification of the normalized distance to the capital is positively and significantly associated with the number of civilian casualties in models 7, 9, 10, and 11. Correspondingly, the smooth estimate in model 12 for the *distance to capital* variable roughly corresponds to this U-shaped form (see figure 5). Again, the strength of the effect is most easily shown in predictions that systematically vary the distance variable while keeping all other variables constant at their means (figure 4).

The control variables for the GED were the same as in the SIGACT analysis for Afghanistan, except for *previous violence,* which was not coded due to the relative scarcity of reported events. Moreover, the number of *military deaths* is coded for each GED event and was included as a control variable. What effects can be expected for the additional control variables? Given the very large share of insurgent events in the SIGACT sample (94.2% of all observations) and the overwhelming qualitative emphasis on the fact that rebels tend to initiate more attacks in insurgencies, one can assume that most of the recorded events in the GED were triggered by insurgents. Therefore, the sample of GED events should mainly consist of insurgent attacks and the directions of the effects should correspond to those of the SIGACT analysis of insurgent indiscriminate violence. As in

the previous analyses, the logged *population* count has a positive estimate, but no significant effect in the full model. A strongly significant and consistently negative effect can be found for the *urban distance* variable which had a positive effect on insurgent indiscriminate violence in the SIGACT analysis. The fact that GED is based on media accounts could drive this result at least partially. Arguably, events in and around large cities are more frequently reported in international media, as the journalistic coverage of rural areas is logistically more difficult. Natural *landcover* is only significant in the baseline model without the main independent variable, but the corresponding estimate is positive, and the results seem too weak to be safely interpretable. The *GECON* variable is again negatively associated with higher levels of indiscriminate violence, which corresponds to the SIGACT models for insurgent activity in Afghanistan. As argued above, military actors tend to fight over economically valuable locations. Multiple violent encounters at those locations are likely to lead to military casualties, but not necessarily to civilian casualties, as civilians can avoid these locations.
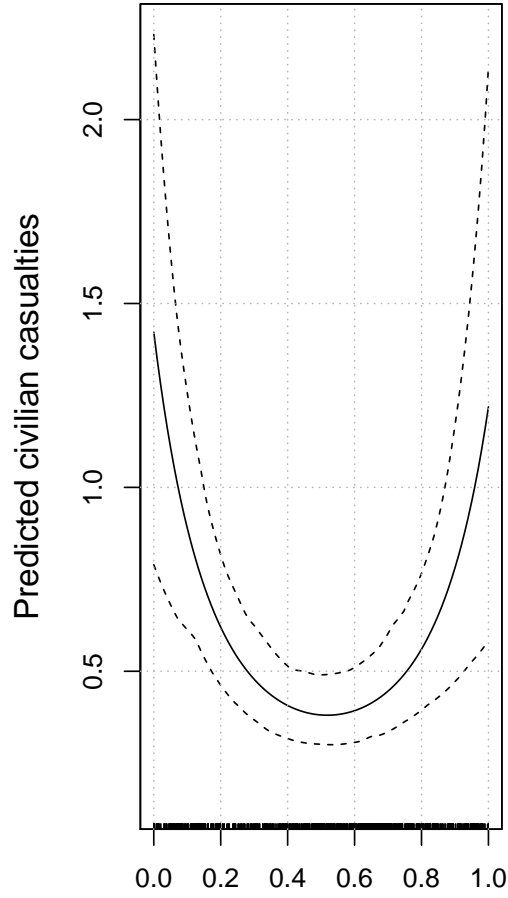
Contrary to the SIGACT analysis, *line-of-sight* is positive and significant for all estimated models. This implies that lowlands are generally more likely to see high levels of civilian casualties than mountainous regions. The high lethality of heavy arms, such as tanks, that cannot be used in forbidding terrain can cause this effect. The *military deaths* variable is also positive and significant for all models estimated for the 10 cases of insurgency, but negative for the unrestricted sample, which is surprising. The positive estimate corresponds to the intuition that higher levels of military casualties also lead to higher levels of civilian victims.

5   Discussion and conclusion

The empirical analysis has confirmed the theoretical expectation of the "Loss of Accuracy Gradient". Focusing centrally on the distance to the incumbent power center, the SIGACT analysis has shown that indiscriminate violence as operationalized through the use of heavy arms is most likely to be applied close to the capital city by insurgents, but far from the capital by incumbent forces. This effect is in line with Hypotheses 1 and 2.

A corresponding analysis with a wider empirical scope has shown that both remote and central regions see the highest levels of civil casualties, which confirms Hypothesis 3. This result holds both for the restricted sample of clearly identified insurgencies, but also for the full GED sample of civil war violence in African conflicts between 1990 and 2010.

Model 7: Civilian Casualties　　　**Model 9: Civilian Casualties**
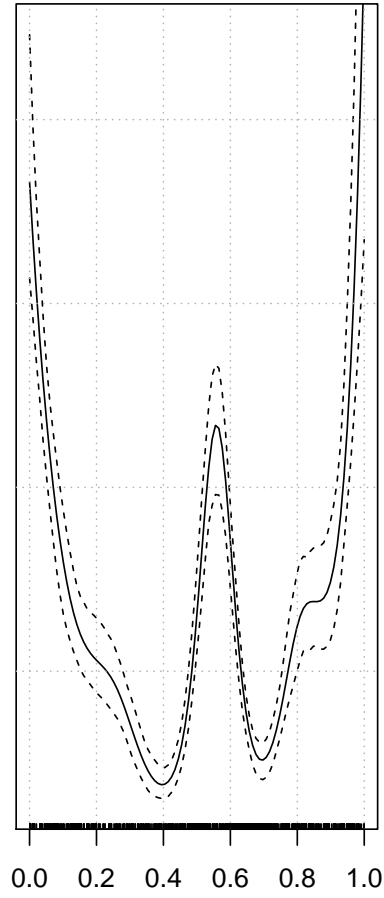
Figure 4: Civilian casualties for cases of insurgency from the GED dataset

Model 9: Smooth estimate for distance to capital

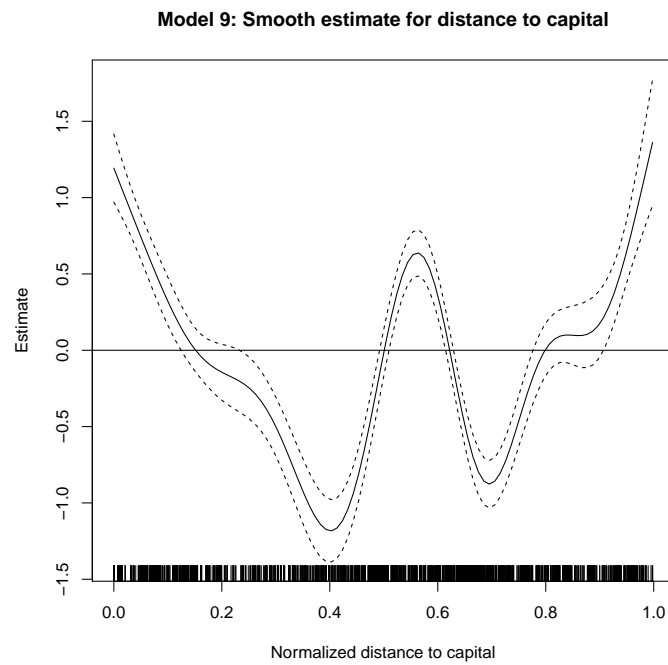Figure 5: Smooth estimate for the effect of *distance to capital* in the GED analysis. Note the U-shaped appearance with a predominantly negative effect in the middle and a positive estimate for very large and very small distances to the capital. The positive estimate around 0.6 could be caused by more intense clashes between the actors at medium distances which also lead to higher civilian casualties.

Most importantly, these results suggest that patterns of indiscriminate violence in civil wars can be predicted ahead of time. Instead of being fully endogenous to levels of military control, initial motivations, competition over resources, or the duration of a conflict, substantial variation in the quality of violence can be explained exogenously. Empirical support for this claim comes from a cross-sectional analysis of a larger number of cases; subsequent research might be able to model areas that are prone to indiscriminate violence more accurately.

These results could therefore serve as a stepping stone towards the more reliable prediction of regions that face a high risk of indiscriminate destruction in civil wars, although it is important to keep in mind that the presented findings are probabilistic tendencies that only partially explain the spatial and temporal distribution of violence in civil wars. Moreover, normalizing the distances to conflict events entails that predictions for future conflicts cannot be easily derived. Nevertheless, this paper has shown that the microdynamics of civil war are substantially influenced by the spatial constellation of the political actors which means that they can be more reliably predicted and modeled than previously thought. This insight is encouraging as it might allow for better informed humanitarian relief operations in the future.

References

Hirotugu Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723, 1974.

Hanna Arendt. *On Violence*. HMH Books, 1970.

Paul Aussaresses. *The Battle of the Casbah*. Enigma Books, 2002.

Ravi Bhavnani, Dan Miodownik, and Hyun Jin Choi. Three two tango: Territorial control and selective violence in israel, the west bank, and gaza. *Journal of Conflict Resolution*, 55(1):133–158, 2011.

Juan Camilo Bohorquez, Sean Gourley, Alexander R. Dixon, Michael Spagat, and Neil F. Johnson. Common ecology quantifies human insurgency. *Nature*, 462:911–914, 2009.

Kenneth Boulding. *Conflict and Defense: A General Theory*. Harper, 1962.

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

Jack E. Bresenham. Algorithm for computer control of a digital plotter. *IBM Systems Journal*, 4(1):25–30, 1965.

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

Halvard Buhaug and Scott Gates. The geography of civil war. *Journal of Peace Research*, 39:417–433, 2002.

Halvard Buhaug and Nils Petter Gleditsch. The death of distance? the globalization of armed conflict. In Miles Kahler and Barbara Walter, editors, *Territoriality and Conflict in an Era of Globalization*, pages 187–216. Cambridge University Press, 2006.

Halvard Buhaug and Jan-Ketil Rød. Local determinants of african civil wars, 1970-2001. *Political Geography*, 25:315–335, 2006.

Halvard Buhaug, Lars-Erik Cederman, and Jan Ketil Rød. Disaggregating ethno-nationalist civil wars: A dyadic test of exclusion theory. *International Organization*, 62(3):531–551, 2008.

Dustin Carpenter, Tova Fuller, and Les Roberts. Wikileaks and iraq body count: the sum of parts may not add up to the whole-a comparison of two tallies of iraqi civilian deaths. *Prehospital and Disaster Medicine*, pages 1–7, 2013.

Lars-Erik Cederman. Articulating the geo-cultural logic of nationalist insurgency. In Stathis Kalyvas, Ian Shapiro, and Tarek Masoud, editors, *Order, Conflict and Violence*. Cambridge University Press, Cambridge, 2008.

Lars-Erik Cederman, Andreas Wimmer, and Brian Min. Why do ethnic groups rebel? new data and analysis. *World Politics*, 62(1):87–119, 2010.

Lars-Erik Cederman, Nils B. Weidmann, and Kristian S. Gleditsch. Inequalities and ethno-nationalist civil war: A global comparison. *American Political Science Review*, 105(3):478–95, 2011.

CIESIN. 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, 2005.

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

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

Sherard Cowper-Coles. *Cables from Kabul*. Harper Press, 2011.

Oliver Crawford. *The Door marked Malaya*. Rupert Hart-Davis, 1958.

DOD. *U.S. Army Counterinsurgency Handbook*. Skyhorse Publishing, 2007.

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

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

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

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

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

David Galula. *Counterinsurgency Warfare: Theory and Practice*. Praeger Security International, 1964.

Dean B. Gesch, Kristine L. Verdin, and Susan K. Greenlee. New land surface digital elevation model covers the earth. *Eos, Transaction, American Geophysical Union*, 80(6):69–70, 1999.

Kristian S. Gleditsch. Transnational dimensions of civil war. *Journal of Peace Research*, 44(3):293–309, 2007.

Kristian S. Gleditsch and Michael Ward. Measuring space: A minimum-distance database and applications to international studies. *Journal of Peace Research*, 38(6):739–758, 2001.

Bernd Greiner. *War without Fronts – The USA in Vietnam*. The Bodely Head, 2009.

Matthew Hansen, Ruth DeFries, John R. G. Townshend, and Rob Sohlberg. Global land cover classification at 1km resolution using a decision tree classifier. *International Journal of Remote Sensing*, 21:1331–1365, 2000.

Håvard Hegre, Gudrun Østby, and Clionadh Raleigh. Poverty and civil war events: A disaggregated study of liberia. *Journal of Conflict Resolution*, 53(4):598–623, 2009.

Kevin Jon Heller. 'one hell of a killing machine': Signature strikes and international law. *Journal of International Criminal Justice*, 11(1):89–119, 2013.

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

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

Macartan Humphreys and Jeremy M. Weinstein. Handling and manhandling civilians in civil war. *American Political Science Review*, 100(3):429–447, 2006.

Stathis Kalyvas. *The Logic of Violence in Civil Wars*. Cambridge University Press, 2006.

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

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

Matthew A. Kocher, Thomas B. Pepinsky, and Stathis Kalyvas. Aerial bombing and counterinsurgency in the vietnam war. *American Journal of Political Science*, 55(2):201–218, 2011.

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

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

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

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

Tse-tung Mao. *On Protracted War*. Foreign Language Press, 1967.

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

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

Claire Metelits. *Inside insurgency*. New York University Press, 2010.

Cerwyn Moore and Paul Tumelty. Foreign fighters and the case of chechnya: A critical assessment. *Studies in Conflict and Terrorism*, 31(5):412–33, 2008.

Andrew Nelson. 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/, 2008.

William Nordhaus, Qazi Azam, David Corderi, Kyle Hood, Nadejda M. Victor, Mukhtar Mohammed, Alexandra Miltner, and Jyldyz Weiss. The g-econ database on gridded output: Methods and data, 2006.

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

Stephen G. Peters, Stephen D. Ludington, Greta J. Orris, David M. Sutphin, James D. Bliss, and James J. Rytuba. Preliminary non-fuel mineral resource assessment of afghanistan. Available online at `http://pubs.usgs.gov/of/2007/1214/`, accessed on July 24, 2013, 2007.

Clionadh Raleigh and Håvard Hegre. Introducing acled: An armed conflict location and event dataset, 2005.

Idean Salehyan. *Rebels without Borders*. Cornell University Press, 2009.

Sebastian Schutte and Karsten Donnay. Matched wake analysis: Finding causal relationships in spatiotemporal event data. *Political Geography*, 41:1–10, 2014.

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

James C. Scott. *The Art of Not being Governed*. Yale University Press, 2009.

Jason Seawright and John Gerring. Case selection techniques in case study research: a menu of qualitative and quantitative options. *Political Research Quarterly*, 61(2), 2008.

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

Ralph Sundberg, Mathilda Lindgren, and Ausra Padskocimaite. Ucdp ged codebook version 1.5-2011. Available online at: http://www.ucdp.uu.se/ged/data.php, 2011.

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

Nils B. Weidmann. On the accuracy of media-based conflict event data, 2014.

Nils B. Weidmann and Idean Salehyan. Violence and ethnic segregation: A computational model applied to baghdad. *International Studies Quarterly*, forthcoming, 2012.

Nils B. Weidmann, Doreen Kuse, and Kristian Skrede Gleditsch. The geography of the international system:the cshapes dataset. *International Interactions*, 36(1):86–106, 2010.

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

Lawrence Wright. *The Looming Tower: Al-Qaeda and the Road to 9/11*. Vintage Books, 2007.

Andrew Zammit-Mangion, Michael Dewar, Visakan Kadirkamanathan, and Guido Sanguinetti. Point process modelling of the afghan war diary. *Proceedings of the National Academy of Sciences*, 2012.

Table 2: Regression results for Indiscriminate Violence in the SIGACT dataset

| | *Dependent variable:* | | | | | |
| | Insurgent indiscriminate violence | | | Incumbent indiscriminate violence | | |
	(1)	(2)	(3)	(4)	(5)	(6)
Distance to Kabul (km)		−0.003*** (0.0002)	−0.003*** (0.0002)		0.002 (0.001)	0.002*** (0.001)
Distance to Pak. (km)		−0.002*** (0.0002)	−0.002*** (0.0002)		−0.004*** (0.001)	−0.004*** (0.001)
Elevation	0.0004*** (0.00003)	−0.0003*** (0.00004)	−0.0003*** (0.00004)	−0.0003** (0.0001)	−0.0002 (0.0002)	
Population	0.00001*** (0.00000)	0.00001*** (0.00000)	0.00001*** (0.00000)	−0.0001 (0.0001)	−0.00004 (0.0001)	
Line-of-sight	−0.002*** (0.0001)	−0.0002 (0.0001)		0.0001 (0.0005)	−0.0002 (0.001)	
GECON	−0.123*** (0.024)	−0.433*** (0.030)	−0.433*** (0.030)	0.546*** (0.102)	0.623*** (0.137)	0.645*** (0.114)
Urban distance	0.001*** (0.0001)	0.001*** (0.0001)	0.001*** (0.0001)	0.0003 (0.0003)	0.0004 (0.0003)	
Landcover	0.156*** (0.011)	0.199*** (0.012)	0.195*** (0.012)	−0.179*** (0.049)	−0.139*** (0.050)	−0.137*** (0.049)
Previous violence	−0.019*** (0.001)	−0.018*** (0.001)	−0.018*** (0.001)	0.017*** (0.005)	0.018*** (0.005)	0.018*** (0.005)
Constant	−2.285*** (0.116)	−0.748*** (0.141)	−0.707*** (0.138)	0.972* (0.504)	0.478 (0.718)	−0.032 (0.419)
Observations	22,512	22,512	22,512	1,288	1,288	1,288
Log likelihood	−13,133.870	−12,739.170	−12,740.190	−813.011	−807.567	−808.711
Akaike Inf. Crit.	26,283.730	25,498.330	25,498.390	1,642.021	1,635.134	1,629.422

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

36

Table 4: Civilian Casualties in the GED Dataset

Dependent variable:

Civilian casualties

	GLM Negative Binomial (link = log)					GAM Negative Binomial (link = log)
	(7)	(8)	(9)	(10)	(11)	(12)
Distance to capital (normalized)	-3.184*** (1.191)		-4.762*** (1.369)	-5.084*** (1.321)	-7.410*** (0.717)	
Distance to capital^2	3.057** (1.309)		4.640*** (1.370)	4.837*** (1.359)	7.359*** (0.700)	
Population (logged)		0.112 (0.071)	0.050 (0.073)			0.120*** (0.027)
Urban distance		-0.003*** (0.001)	-0.003*** (0.001)	-0.003*** (0.001)	-0.001*** (0.0002)	-0.002*** (0.0003)
Landcover		0.050 (0.036)	0.004 (0.037)			0.054*** (0.012)
GECON		-7.506*** (1.858)	-7.858*** (1.901)	-7.387*** (1.728)	-5.948*** (1.177)	-14.937*** (0.951)
Elevation		0.001*** (0.0002)	0.001*** (0.0002)	0.001*** (0.0002)	0.0005*** (0.0001)	0.001*** (0.0001)
Line-of-sight		0.001*** (0.0005)	0.002*** (0.0005)	0.002*** (0.0004)	0.001*** (0.0002)	0.001*** (0.0001)
Military casualties		0.003 (0.002)	0.006*** (0.002)	0.006*** (0.002)	-0.001* (0.0004)	0.004*** (0.0004)
Constant	0.353 (0.227)	-2.144*** (0.673)	-0.644 (0.865)	-0.132 (0.331)	0.502** (0.203)	-2.132*** (0.263)
Observations	3,649	3,649	3,649	3,649	10,406	3,649
Adjusted R^2						-496.323
Log Likelihood	-2,153.948	-2,125.379	-2,119.689	-2,119.842	-6,927.447	-4,031.412
θ	0.030*** (0.002)	0.034*** (0.002)	0.034*** (0.002)	0.034*** (0.002)	0.037*** (0.001)	
Akaike Inf. Crit.	4,313.896	4,266.759	4,259.378	4,255.684	13,870.890	

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