The strategic logic of ethno-territorial competition.

Violence against civilians in Africa’s civil wars

by

Andreas Wimmer and Chris Miner

forthcoming in the *Journal of Global Security Studies* 4, 2019

Abstract:

This article empirically confirms a new theory of violence against civilians during civil wars using geo-coded data on African armed conflicts over the past two decades, combined with a range of other geo-coded information. The theory outlines a twofold logic of ethnic targeting. First, rebels and government fighters kill civilians in areas populated in equal shares by their own and their adversary’s co-ethnics because in such areas, small amounts of violence suffice to tilt the local balance of power in their favor. Second, they target places close to the border between the settlement areas of their own and their adversary’s co-ethnics. Both strategies aim at establishing political control over ethnic settlement areas where this is strategically most useful and most effective. In contrast to ethnic targeting, the three most prominent theories assume that civilian victimization is independent of the political conflict associated with the civil war. We don’t find empirical support for these alternative theories: Civilians are not more likely to be killed the closer they live to lootable natural resources, in recently conquered territories where fighters are supposed to eliminate enemy collaborators, or where rebel forces operate that have established only weak control over their fighters.

In the wake of Kalyvas’ (2006) seminal study of the Greek civil war, scholars started to more systematically investigate why armed organizations intentionally kill unarmed civilians during the course of a civil war. This paper makes a three-fold contribution to this literature. First, it

---

1 The author would like to thank audiences at the workshop “Bridging Micro and Macro Approaches to Civil War and Political Violence” organized in Barcelona by Laia Balcells, at the Graduate Institute in Geneva, and at the department of sociology at Princeton University.
joins other recent work in providing evidence that the killing of civilians is linked to their ethnic background, and it specifies the precise logic of ethnic targeting. Previous case studies have shown that ethnic affiliations matter for understanding who kills whom.\(^2\) Going beyond these individual cases, this article looks at the larger universe of African conflicts and shows that the logic of ethnic targeting is important for understanding where and when civilians are killed during civil war, even when other mechanisms, highlighted by other theories, are taken into account and even if we look at the entire empirical universe that includes ethnic as well as non-ethnic armed conflicts.

Second, this article theoretically disentangles and empirically evaluates two distinct theoretical models of why and how ethnic targeting occurs. Fjelde and Hultman (2014) build on Valentino and co-authors (2004) to argue that armed groups attack their adversary’s co-ethnic civilians in order to undermine its logistic and informational support base and thus gain an advantage in winning the civil war. Armed actors kill civilians, in other words, to undermine the enemy’s fighting capacities. They therefore should attack their opponent’s co-ethnics where these are more numerous and thus easier to find and target.

We outline an alternative logic of ethnic targeting, knitting together various existing arguments into an integrated theoretical framework. Violence is about expanding the area under the political control of one’s co-ethnics, we argue, and is thus most intense where it is strategically most useful and most effective. Rather than weakening the enemy’s civilian supporters where they are most numerous, violence is therefore more precisely and strategically focused on specific areas. Border zones between their own and their adversaries’ co-ethnics are of greater strategic interest than ethno-territorial enclaves because fighters are interested in expanding the continuous

\(^2\) For Darfur, Olsson and Siba 2009; for Aceh, Czaika and Kis-Katos 2009; for Bosnia, Costalli and Moro 2010; Weidmann 2011; for Guatemala, Sullivan 2012; Schwartz and Straus 2018; for Northern Ireland, White 1993.
territory controlled by supportive civilians (similarly Melander 2007; Weidmann 2011). Second, in locales where opposed groups make up equal shares of the local population, small amounts of violence can tilt the local demographic and political balance of power in one’s favor by intimidating the opposed group and encouraging them to flee, thus expanding the territory under the political control of co-ethnics in the most cost effective way. This second argument is derived from Balcells (2011) study of the civil war in Spain.3 Killing civilians in order to expand the territory controlled by co-ethnics may be part of a broader and more radical strategy of ethnic cleansing: an attempt to create ethnically homogenous regions or states by discriminating against or terrorizing members of ethnic minorities such that they would flee or by even actively organizing their deportation (Petrovic 1994).4 Unfortunately, we cannot determine whether ethnic targeting is part of a broader strategy of ethnic cleansing because there is no systematic data on the war goals and strategic visions of armed actors.

The third contribution of this article is to evaluate three alternative theories of violence against civilians as well, using varying datasets, samples, and definitions of the dependent variable to mirror as closely as possible the specific scope conditions and mechanisms foreseen by each of the three theories. In contrast to the model of ethnic targeting outlined above, these three alternative theories assume that killing civilians has little to do with ethnicity or, more generally, with the macro-level political conflict over which a civil war is fought. Greed theories maintain that the population in resource rich areas will be terrorized and victimized by armed groups who seek to establish or maintain control over these resources (Azam and Hoeffler 2002; Bagozzi 2017; Eyndre 2016; Koren and Bagozzi 2017). According to Weinstein (2006; see also

---

3 It is also impossible to explore the symbolic and emotional drivers, such as resentment, fear, and mythologies demonizing the ethnic enemy, that provide rank-and-file fighters with the motivations for killing civilians (see Kaufman 2006; Petersen 2002).

4 Mann (2005), however, uses the term synonymously with genocide aiming at the physical destruction of a group.
Humphreys and Weinstein 2006), rebel organizations that initially relied on funding from natural resource extraction or from outside governments attract opportunistically motivated fighters who will be more likely to prey upon the civilian population (except if the outside government is a democracy, see Salehyan et al. 2014). A second theory was proposed by Kalyvas (2006). Rebels and governments act upon the denunciations of local villagers, he argues, to kill possible collaborators. They are most likely to do so when they don’t yet fully control an area and thus are already able to search for and eliminate collaborators but haven’t killed all of them yet. Seen from these various perspectives, ethnicity does not define friends and foes in clear-cut terms, and “ethnic defections” are therefore thought to be frequent: Armed organization whose members share the same ethnic background sometimes fight each other (Christia 2008) or even kill co-ethnic civilians (see Kalyvas 2008; Lyall 2010).

Testing these various arguments, this article improves over other cross-national studies that either focus on one of them (Wood 2010; Wood et al. 2012; Fjelde and Hultman 2014; Salehyan et al. 2014) or test other propositions that are only tangentially related to these most widely discussed theories.⁵ To do so, we use a range of geo-coded datasets, different samples of areas and

⁵ A booming series of recent research explores additional, more fine grained arguments about characteristics of violence-prone rebel organizations, the conditions under which they resort to indiscriminate violence, and the effects of outside intervention. More precisely, recent scholarship found that democratically oriented rebel organizations commit fewer acts of violence than religiously oriented organizations in the Middle East (Asal et al. 2015), that Communist organizations rape women less than other organizations (Hoover Green 2016), that more civilians are killed by a rebel force after a competing rebel organization has emerged (Wood and Kathman 2015) or after the rebels have sustained many casualties at the hand of government forces (in Uganda and Africa more generally: Wood 2014). Somewhat contra-intuitively, one-sided violence against civilians is more frequent in conventional civil wars than in guerilla insurgencies (Krcmaric 2018). Regarding the effects of outside intervention, Hultman and Kathman (2013) have shown that UN peacekeeping missions reduce civilian victimization in Africa’s civil wars while Wood and co-authors (Wood et al. 2012) find that armed intervention by outsiders decreases the killing of civilians by the supported civil war actor and increases victimization by the opponent. Finally, more civilians are killed when many
conflicts, and different, perpetrator-specific definitions of the dependent variable, seeking to evaluate these theories with the most appropriate data mirroring the scope conditions they specify. The dependent variables are derived from either the Armed Conflict Location and Events Dataset (ACLED; Raleigh et al. 2010) or the Uppsala Conflict Data Program’s Geo-Referenced Event Dataset (UCDP-GED; Sundberg et al. 2012) that cover the African continent from 1997 to 2011 and from 1989 to 2010 respectively. In combination with a variety of other geo-coded datasets, most importantly the Ethnic Power Relations dataset (Geo-EPR; see Wucherpfennig et al. 2011), these data allow me to identify which armed actor intentionally killed civilians in which exact location, what the ethnic composition of that location is, and with which ethnic populations armed actors themselves identify.

We specify logistic regressions with 0.25-degree hexagons (approx. 27.75 km²) as units of observation and violence as the dichotomous dependent variable. Fixed effects control for unobserved characteristics at the country level, such as the nature of the conflict, levels of rule of law, and the like. We find strong and consistent support for the ethno-territorial competition argument: violence against civilians is most likely in areas where co-ethnics of rebels and government elites reach demographic parity as well as close to the border between the settlement areas of these two groups. Both findings indicate that expanding the territory dominated by one’s co-ethnics is an important strategic aim of killing civilians.⁶ we don’t find support for the idea—developed empirically by Fjelde and Hultman (2014)—that ethnic targeting aims indiscriminately at the opponent’s co-ethnics where they are most numerous and thus easiest to find and target.

---

⁶ Since the UCDP-GED data on civilian victims are limited to Africa, I am unable to evaluate whether these results would hold for Latin America, Europe, or Asia but note that past research on Bosnia (Weidmann 2011) and Spain (Balcells 2011) has led to similar conclusions.
Online Appendix 3 outlines why Fjelde and Hultman, who use similar data sources and methods, arrived at different results, which we briefly summarize in the main text.

The results also don’t support the looting, loose control over fighters, and territorial control arguments, or even offers evidence that contradicts them: civilians are not more likely to be victims of violence if they live close to lootable diamond fields, to other mining sites, or to oil wells. Weak control over rebel fighters or funding from natural resource extraction decreases—rather than increases—their propensity to kill civilians, while rebels’ external funding is not associated with civilian victimization. Both government troops and rebels are more likely to kill civilians close to where they lost territory to their opponents (in line with Wood 2014), rather then where they gained it, as Kalyvas’ model implies.

The article proceeds in a straightforward way. The following section discusses the various theories in more detail and derives empirically testable hypotheses from them. Next we introduce datasets and units of observation, define independent and dependent variables and discuss model specifications. A section with results follows and a final section concludes.

**Arguments and hypotheses**

*The ethno-political logics of violence*

This article joins a group of authors (Balcells 2011; Valentino et al. 2004; Hirose et al. 2017; Schwartz and Straus 2018) who see violence against civilians as part and parcel of a broader conflict over political power, rather than a purely local issue of controlling rebel fighters, natural resources or containing collaboration, as in the three most prominent arguments in the literature. In this article, we focus on conflicts were political goals and adversaries are defined in ethnic
terms, rather than with reference to class, party, or ideology, mostly because data on party and class affiliation or the ideological orientation of the civilian population is hard to come by (for an example of what it takes to overcome these challenges, see the Afghan study by Hirose et al. 2017). We do not imply that there must be something in the nature of the ethnic bond that makes it more likely that civilians are killed, nor that Africa is more prone to ethnically defined conflicts than other regions, nor do we think that ethnic conflicts are generally more violence-prone than other types of conflicts (a view disproven by Valentino et al. 2004).

we assume that ethnic conflicts are fought over the distribution of political power among ethnically defined alliance networks (in line with Wimmer 1997; Wimmer 2013). The following figure describes various such configurations of ethno-political power, modifying Tilly’s (1978) well-known polity model. It distinguishes between ethno-political groups represented at the highest level of government (termed “included” groups A to C) and those that are not (“excluded groups” 1 to 4). Two different types of conflict can be distinguished. A rebel organization claims to speak in the name of excluded group 1 and fights central government e.g. over the neglect of its home region or the lack of representation in the inner circles of power. We call this type of conflict a “rebellion” (see Wimmer et al. 2009). A well-known example would be the various Darfurian guerillas that fought against an Arab dominated Sudanese government.

Second, an army faction of group A members (to give an example) attempts to overthrow the existing government, claiming that the latter has delayed the promotion of A-officers in the army and disempowered the ministries controlled by A. We call this type of conflict “infighting” (ibid.). An empirical example are the Diola of Senegal, representatives of whom are included in the power sharing coalition since independence; from the mid-nineties onward, a separatist organization, largely dominated by Diola speakers, began to fight for the independence of the
Casamançé region. On the basis of these analytical distinctions, we can now outline two different theories that predict which actors will target which civilian populations during an armed conflict.

Figure 1
Ethno-political actors and conflicts

Ethno-territorial competition

According the theory proposed here, fighters seek to maximize the territory under control of their co-ethnics and they do so in a strategic way. Before we outline this argument in detail, we first clarify why ethnic ties are politically relevant and how local struggles for political power are intertwined with the macro-level conflict. The local population (group 1 in Figure 1) gains from supporting ethnic rebels who fight in their name because this increases the likelihood that they will be able to win the conflict and capture the state. This in turn will produce patronage and public goods benefits for the allied ethnic population (for evidence of ethnic favoritism in public goods provision see Bannjerjee et al. 2008, section 3.2.1; de Luca et al. 2015; Grødeland et al.)
2000; McClendon 2016; Kramon and Posner 2012). For the same reasons, the co-ethnics of ruling elites (groups A, B, and C in figure 1) have an interest in maintaining their privileged relationship with the state apparatus and thus are more likely to support government troops than excluded populations.

These national struggles over state power are then mirrored at the local level, where political factions and networks align with the ethno-political cleavage at the macro level, thus opposing representatives of group 1 against those of groups A, B, and C. Consequently local struggles over political power—over “who rules this place”—become intertwined with national level conflict over “who owns the state” (Wimmer 1997).

In this local context, fighting organizations kill civilians of the other side in order to tilt the balance of local political power and demographic weight in the favor of their own group, intimidating the enemy population or encouraging them to flee. They do so because they can count on the political loyalty and logistical support of their co-ethnics (cf. also Wood 2010: 602f.). The marginal utility of violence is highest in locales where it would suffice to expel (or kill) only a small number of civilians in order to become a local majority. Therefore, violence should be most likely in those areas where the shares of the co-ethnics of government elites and of rebels reach parity (H1).

For government violence, this means maximum polarization between groups 1 and ABC or between A and BC, for rebel violence between groups 1 and ABC, and for violence committed by infighters between A and BC. Conversely, where the co-ethnics of fighters represent 90% of the population, it makes little sense to act violently against what most likely already is a quiescent and subdued minority. Where co-ethnics make-up only 10% of the population, the amount of violence needed to expel at least 40% of the other group is prohibitive—and it might even be undesirable if most people perceive these locales to represent the “homeland” of their enemies.
A similar argument was first developed in Balcell’s (2010; 2011) study of the Spanish civil war fought in the Catalan and Aragon regions. We integrate it into my model of ethno-territorial competition.\(^7\) we also go one step beyond Balcell by taking a second strategic interest of fighting organizations into account as well. Armed groups are interested in establishing control over regions where the settlement areas of their co-ethnics and those of their opponents meet each other, in other words, in border areas (a more fine-grained version of this argument has been proposed for Bosnia by Melander 2007; Weidmann 2011). The strategic utility of a co-ethnic village in a border zone is much higher than that of an otherwise similar co-ethnic exclave deep in territory populated by one’s opponent’s co-ethnics. Thus, violence should be more intense the closer a locale lies to a border between the co-ethnics of rebels and government elites (H2; the border between 1 and ABC or A and BC for government violence, 1 and ABC for rebel violence, and A and BC for infighters).\(^8\) These two hypotheses are markedly different from those derived from the “weakening the enemy” perspective, to which we now turn.

*Weakening the enemy’s ethnic support base*

This second theory of ethnic targeting has been proposed by Fjelde and Hultman (2014). Armed groups attack co-ethnics of their enemy, they argue, in order to undercut its logistical, informational, and material support. This makes strategic sense since in ethnic conflicts, the ethnic background of a civilian provides a good informational shortcut to determine her political

---

\(^7\) Weidman (2011) confirmed the polarization hypothesis with regard to Bosnia, but argued that it emerged because of local competition for jobs, public goods, and other local resources. For a critique of explaining ethnic conflicts with economic competition arguments, see Horowitz 1985:105-135; Author Citation.

\(^8\) Border areas are also those where most polarized locales can be found (average polarization index of .28 compared to .0007 in hexagons borders).
leanings (ibid.; Olsson and Siba 2009; Sullivan 2012). In the set-up described by Figure 1, government troops should therefore attack civilians of groups 1 and A (H3) while sparing civilians of groups BC. Rebels associated with group 1 in turn should attack all included groups ABC (H4) and avoid groups 1 to 4 (and especially 1), while infighters from group A should attack groups BC (H5) and avoid all others (especially their own group A). Since there are no further strategic considerations according to this theory, rebels and government troops should kill civilians of these target groups wherever they live. By implication, violence should be more intense the higher the share of the target population in an area (which is roughly what Gulden [2001] finds for Guatemala).

Fjelde and Hultman, using similar data sources as in this project, test whether areas that contain at least some co-ethnics of the adversary see more violence than areas without any of the adversary’s co-ethnics.9 However, this coding of the independent variable is not capturing the supposed mechanism in the best way. If fighters seek to weaken their opponents’ support base, they should kill more civilians the larger the local share of their opponents’ ethnic base in an area. We therefore test their theory with a dummy variable for hexagons entirely populated by the adversary’s co-ethnics. Alternatively, we used a continuous measurement (the percentage of the enemy’s co-ethnic population), which produced substantially similar results (for details see Online Appendix 3).

Other theories: Controlling resources, fighters, or territory

We now briefly discuss the three major theories of civilian victimization, all of which claim that where and why civilians are killed has nothing to do with the political struggle between

9 Using percentage of enemy co-ethnics, rather than a dummy variable, Sullivan (2012) and Olsson and Siba (2009) arrive at similar conclusions for Guatemala and Darfur.
insurgents and governments and is thus entirely independent of civilians’ stance in that macro-political conflict. In other words, violence against civilians is supposed to follow a non-political logic, determined by either economic considerations (gaining access to natural resources to enrich the fighters and/or fund their operations), organizational logics (whether or not rebel leaders can keep their fighters under control), or locally specific military tactics (eliminate possible collaborators in an area).

Looting: The looting argument derives from the greed school in the civil war literature (Collier and Hoeffler 2004; Le Billon 2001), according to which governments and rebels fight to gain control over lootable resources, such as the famed “blood diamonds” of Sierra Leone, oil resources in the Nigerian delta, or the rare metals found in Eastern Congo. Azam (2002) has applied this mode of reasoning to violence against civilians (see also Querido 2009; Hegre et al. 2009). In his game-theoretic model, violence against civilians represents a side effect of the struggle over lootable goods, which are needed to pay fighters. Once violence against civilians has set in, young men have incentives to join an army or rebel force rather than to farm their land, leading to a self-reinforcing equilibrium at a high level of violence. The hypothesis derived from this perspective is straightforward: Violence against civilians should be more frequent in areas with abundant lootable natural resources, including diamonds, oil, and minerals (H6).

Loose control over fighters: Weinstein (2006) focuses on the initial conditions that shape a rebel organization and thus its subsequent behavior vis-à-vis the civilian population. Rebel

---

10 Other authors have specified the looting argument differently and maintained that civilians are killed in areas with rich crop harvests (Koren and Bagozzi 2017), especially during periods of draught (Bagozzi 2017; for India Eyndre 2016), or where aid organizations (Wood 2015) or national governments (Khanna and Zimmerman 2017; for Columbia: Weintraub 2016) provide locals with lootable resources. In order not to complicate the picture and for reasons of data availability, I don’t attempt to test these additional greed arguments but note here, en passant, that in the statistical models discussed below, violence is not less likely in places with sparse vegetation or where the soil quality is bad and thus less suitable for agriculture.
organizations that were funded either by natural resources or by external actors such as a foreign
government attracted more opportunistic fighters interested in immediate economic gains rather
than in achieving a political goal. Such organizations will also exhibit a lower level of control
over the rank-and-file and less hierarchical integration. As a consequence, these rebel forces are
more likely to prey upon the local population and kill civilians (for Sierra Leone see Humphreys
and Weinstein 2006). Three hypotheses are derived from this approach: Rebel organizations that
command less control over their fighters (H7) or have relied on natural resources to fund their
operations (H8) or on external financial support\textsuperscript{11} (H9) should be more likely to commit violence
against the civilian population.

\textit{Containing collaboration:} According to Kalyvas’ (2006) well-known study of the Greek civil
war, rebels and governments in irregular civil wars selectively kill civilians in order to eliminate
collaborators and prevent future defections. This logic operates independently of the macro-
political conflict around which the civil war is fought. Rebel and government soldiers kill
civilians when they haven’t yet fully established control over a territory. They are less likely to do
so when they have either no control (in contested areas) or already fully established themselves
and thus eliminated any opposition to their rule. Civilians, on the other hand, denunciate each
other to settle scores from local level feuds again unrelated to the macro-conflict. This is most
risky if the area is heavily contested between war participants where civilians do not know who
will eventually prevail. The combination of these two behavioral logics leads Kalyvas to expect
that violence is highest where a good supply of denunciations meets a high demand for
identifying collaborators, that is, in areas of intermediate control. Such intermediate levels of

\textsuperscript{11} In an extension of the argument, Salehyan and co-authors (Salehyan \textit{et al.} 2014) show at the level of rebel
organizations that violence against civilians decreases if an organization is externally supported by a democratic
regime, and increases if the supporting state is not-democratic. For data reasons I again refrain from testing this more
nuanced version of the argument.
control should be found in areas that recently shifted from one side to the other (H10) without, however, having shifted hands many times in the recent past, which would indicate a contested area that the theory expects to be more peaceful.

The set of theories discussed so far represent the most widely cited and empirically specific approaches to violence against civilians in civil wars. Not all of them are incompatible with each other or mutually exclusive: For example, it is well possible that a looting logic combines with that of ethno-territorial competition. This article does not attempt to discover such more complex causal relationships, but seeks to evaluate whether some of the basic observable implications of each theory hold empirically. Future work could certainly do more to discern possible interaction effects or to identify the scope conditions under which one or the other of these theories is more likely to hold, even if a general analysis does not produce statistically significant results in their support. The next section discusses the data and variables used to evaluate the various hypotheses introduced above.

**Data and model specification**

*Dependent variables and units of observation*

As indicated before, there are two available datasets with geo-coded information on violence against civilians (ACLED and UCDP GED). For reasons discussed below, we use the UCDP dataset for all analysis except to evaluate Kalyvas’ theory, for which the ACLED dataset is more appropriate. The dependent variable in both datasets is defined as an event during which violence against unarmed civilians was committed intentionally by government soldiers, rebel forces, or militias—thus ignoring “collateral damage” of armed battles. Data on violent events is much
more reliable than estimates of the number of deaths. We therefore define the dichotomous dependent variable as any event during which an armed group intentionally killed at least one civilian. For robustness purposes, we ran the main models with a count variable of the number of civilian deaths and report the results in Table 3.

The UCDP GED dataset is limited to Africa. It covers all years from 1989 to 2010 and all countries where armed groups killed at least 25 civilians. It can be linked to other datasets on armed conflict, conflict participants, their external support, their links to politicized ethnic groups, etc. Since all the theories discussed above relate to violence during civil war, we limit the universe of observation to years and countries during which an armed conflict was active, i.e. where at least 25 battle deaths were counted in the UCDP armed conflict dataset (Gleditsch et al. 2002).

The ACLED dataset includes some conflict prone countries in the Middle East and Asia as well. It has been criticized for a number of reasons, most importantly related to the quality of the coding (Eck 2012; the recent updates by ACLED seem to have addressed these issues). Compared to UCDP, ACLED has the advantage of offering information on territorial gains and losses by rebels and government troops, which allows evaluating Kalyvas’ theory. For the rest of the analysis, however, we rely on UCDP data because, as mentioned above, it is linked with other datasets relevant for this study.

Both datasets identify perpetrators and we can therefore distinguish between violence committed by government or rebel organizations. In order to test the two ethnic targeting arguments, we

---

12 I prefer this dichotomous coding of the dependent variable over a count version since we are not theoretically interested in how many violent episodes a hexagon experiences, but in the more fundamental contrast between peace and violence. This also represents the explanatory focus of the theories we intend to evaluate. Moreover, some of the reliability problems in event counts based on media reports (see Weidmann 2012)—such as how to distinguish separate episodes from each other—are reduced when using a dichotomously coded variable.
need to further disaggregate the dependent variable and distinguish between violence committed by infighters and by rebels, and we need to identify the ethnic communities that each armed organization claims to represent. This is made possible thanks to the actor identification number that is offered both in the UCPD-GED and the EPR datasets (Wucherpfennig et al. 2012).

Both the ACLED and the UCDP datasets contain precise geo-coding for events. In order to make this data usable for a cross-sectional time-series analysis, we overlaid a grid of fixed territorial units in the form of hexagons of 27.75-km2 size, or one quarter of a degree of latitude and longitude. We chose hexagons as units of observation, rather than settlement units such as villages and towns, because much of Africa’s population South of the Sahara lives in sparsely populated and dispersed hamlets outside of towns and villages (Herbst 2000). More importantly, there are no data on each hamlet or village’s ethnic composition and we thus have to rely on the coarser ethnic maps provided by the GeoEPR dataset (see below). The robustness section reports results from other specifications of the units of observation, such as larger rectangles or politically defined units of various size.

we disregarded the fine-grained day-by-day coding of violent events in both datasets and instead aggregated to the year—the level of granularity of many other independent variables. For exploring Kalyvas’ argument, however, a yearly data resolution might not be fine-grained enough. It may well be, for example, that government troops gain control of a territory in December and kill civilians during January—a link that will be lost in a dataset based on years.

We therefore created a dataset with month-hexagons as units of observations, using the ACLED dataset for the reasons discussed above, and we then worked with one-, two-, and three-month lags of the independent variables.
Independent variables

The following gives a short overview of the various independent variables coded for this project. We refer the reader to Online Appendix 1 for more detail on coding procedures and data sources. Online Appendix 2 provides summary statistics. To assess greed theories of violence against civilians, we focus on three different types of lootable resources: the number of active, on-shore oil wells (from Lujala et al. 2005); whether or not there are diamond deposits in a hexagon (data from Gilmore et al. 2005); the number of economically significant mining sites (e.g. for Tungsten and other rare metals) per hexagon. Other codings of these variables (e.g. the km distance to natural resource sites) were used for robustness purposes and produced similar results.

To evaluate Weinstein’s argument, we rely on Cunningham and co-authors’ (2009) dataset on rebel organizations to code a dichotomous variable indicating whether or not a rebel organization received any kind of financial support from actors situated outside of the focal country. The second variable measures the strength of the control that rebel leaders exercise over the rank-and-file. From another dataset by Rustad and co-author (2012), we use the information whether or not a rebel organization has been funded by natural resources (including drugs). As with the external financing variable, there is no information on whether rebels relied on such resources when they first organized, or later on, or both—we thus cannot test Weinstein’s path dependency argument in a fully adequate way. To avoid confounding civilians killed by other rebel organizations with different internal organization or funding sources, we restrict the sample, for these statistical models, to the areas within which a rebel organization operates.

Kalyvas’ theory is the hardest to evaluate empirically because there is no data on the actual levels of control over territory by specific actors. It is perhaps not too far-fetched, however, to assume that in recently acquired territories, an armed group has neither established full control yet nor continues to share control with its adversary. Data on territorial gains and losses is available in
ACLED, as discussed above. We created two count variables indicating how many times either government or rebels have gained territory on a hexagon, whether through battle or peacefully. To check for robustness, he also coded a distance variable—assuming that the degree of control over territory monotonically decreases with distance to the location of a territory change. The results are substantially very similar.

To exclude the possibility that we are looking at territories that frequently switch back and forth between rebels and government troops (where civilians should be safer according to Kalyvas), we excluded hexagons with territorial gains in favor of both rebels and governments in the previous 12 months. To take two additional scope conditions of Kalyvas’ theory into account, we excluded conventional civil wars from consideration (using the war typology in Kalyvas and Balcells 2010), since the theory applies to guerilla wars only. This reduced the number of events from 2834 to 2506. We also excluded events with more than 5 civilian deaths from the coding of the dependent variable because indiscriminate killings (such as massacres) are also not the focus of Kalyvas’ argument (we adopt the threshold to define indiscriminate violence from Sullivan 2012). This meant that we excluded an additional 983 events. For robustness purposes, we defined the dependent variable in a less restricted way to include contested territories, guerilla wars, and high levels of violence with more than 5 civilian victims. The results, shown in Table 3, are similar.

To evaluate the two ethnic targeting arguments, we rely on the Ethnic Power Relations (EPR) dataset (Wimmer et al. 2009), which also exists in a geo-coded version for the subset of ethnic groups that have identifiable territorial settlement patterns (Wucherpfennig et al. 2011; we use the EPR-ETH edition of the data). The EPR dataset identifies all politically relevant ethnic
groups\textsuperscript{13} and their position in the overall power configuration rendered in Figure 1. EPR also codes whether a rebel organization has emerged that fights in the name of a particular ethnic group, providing a rebel identification number that links EPR to other datasets (Wucherpfennig et al. 2012).

This allows identifying the local, hexagon-specific population shares needed to evaluate the two ethnic targeting arguments: the share of co-ethnics of government elites (ABC in Figure 1), of rebels (1), of infighters (A), of infighters’ adversaries (BC), as well as the corresponding polarization figures. We use Montalvo and Reynal-Querol’s (2005) well-known formula to calculate the latter. Finally, Geo-EPR allows to calculate the kilometer distance to the nearest border between the settlement territories of groups in conflict, i.e. the co-ethnics of rebels or infighters on the one hand and government’s co-ethnics on the other hand (for details see the Online Appendix 1). Because the logic of ethnic targeting should not be observed in areas where ethnicity is not a relevant factor for political competition, we restrict the sample to areas

\textsuperscript{13} To be more precise, an ethnic category is politically relevant if at least one actor (a political movement, or a party, or an individual) with some minimal resonance in the national political arena claims to speak for this category or if outsiders consider the category to be relevant by discriminating against group members (see for the following, Wimmer et al. 2009). EPR does not code individuals, parties, or movements as “representing” an ethnic group if these actors cannot acknowledge their ethnic background in public or publicly pursue group interests. In line with constructivist notions of ethnicity, the list of relevant categories can change over time, and categories can fission or fuse. EPR is based on an encompassing definition of ethnicity (see, e.g., Wimmer 2008) that includes groups with a distinctive religion, language, race, culture, or profession (as in caste systems).

EPR lists the political status of each ethnic category for each year by evaluating whether group members can be found at the highest levels of executive government, such as the cabinet in parliamentary democracies, the ruling circle of generals in military dictatorships, the politburo in communist countries, and so on. The measurement is thus independent from whether the regime is democratic or not. Levels of representation are measured through an ordinal scale. It ranges from monopoly power (total control of executive government by representatives of a particular group) to dominance (some members of other groups hold government positions), senior partner in a power-sharing arrangement, junior partner, representation at the regional level (e.g., in a provincial government), powerless, and discriminated against (i.e., targeted exclusion from any level of representation).
populated by at least one politically relevant ethnic group, but loosen this restriction when running robustness models.

The probability that armed groups kill civilians depends on other, theoretically less interesting factors that relate to the sheer number and accessibility of civilians. We experimented with a large number of variables from many different geo-coded datasets and produced a series of baseline models that were then tested for multi-collinearity problems. Online Appendix 1 describes these variables and the data sources.

To facilitate orientation, the Table 1 gives an overview of the main arguments, hypotheses, the operational measurements, the datasets and their temporal resolution, as well as the exact specifications of the dependent variable.

Table 1: Arguments, hypotheses, and measurements

<table>
<thead>
<tr>
<th>Main argument</th>
<th>Looting</th>
<th>Loose control</th>
<th>Containing collaboration</th>
<th>Weakening the enemy</th>
<th>Ethno-territorial competition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Civilians who are in the way of accessing and exploiting lootable resources will be killed</td>
<td>Rebels relying on external aid or natural resource funding will have less control over rank-and-file and will kill more civilians</td>
<td>Civilians are killed where rebels or government soldiers haven’t yet established full control</td>
<td>Civilians are killed if they share ethnic ties with the attacker’s opponent</td>
<td>Civilians are killed in localities close to settlement borders between groups in conflict and where these reach parity</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Main motivation for killing civilians</th>
<th>Control over resources</th>
<th>Looting by rank-and-file</th>
<th>Eliminating collaborators</th>
<th>Weaken civilian support of enemy</th>
<th>Expanding area under control of co-ethnics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control over resources</td>
<td>Looting by rank-and-file</td>
<td>Eliminating collaborators</td>
<td>Weaken civilian support of enemy</td>
<td>Expanding area under control of co-ethnics</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Do political affiliations of victims matter?</th>
<th>No</th>
<th>No</th>
<th>No</th>
<th>Yes</th>
<th>Yes</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Does ethnic background of victims matter?</th>
<th>No</th>
<th>No</th>
<th>No</th>
<th>Yes</th>
<th>Yes</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Where the theory predicts civilian victims</th>
<th>1) Areas with many oil producing wells 2) Areas with diamond deposits</th>
<th>1) Area with rebels that receive outside financial support 2) Areas with</th>
<th>1) Areas in which government acquired territory are targeted by</th>
<th>1) Areas populated by the co-ethnics of rebels are targeted by government</th>
<th>1) Areas close to the settlement border between groups in conflict</th>
</tr>
</thead>
<tbody>
<tr>
<td>3) Areas with many economically significant mining sites</td>
<td>rebels that are funded by natural resources</td>
<td>government troops in which rebels acquired territory are targeted by rebels</td>
<td>troops</td>
<td>2) Areas populated by co-ethnics of governing elites are targeted by rebels</td>
<td>2) Areas with high levels of polarization between co-ethnics of governing elites and of rebels</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Dataset to code dependent variable</td>
<td>UCDP</td>
<td>UCDP</td>
<td>ACLED</td>
<td>UCDP</td>
<td>UCDP</td>
</tr>
<tr>
<td>Temporal resolution</td>
<td>Yearly</td>
<td>Yearly</td>
<td>Monthly</td>
<td>Yearly</td>
<td>Yearly</td>
</tr>
<tr>
<td>Sample restrictions</td>
<td>Only countries and years with an ongoing civil war</td>
<td>Only countries and years with an ongoing civil war</td>
<td>Only countries and years with an ongoing civil war</td>
<td>Only countries and years with an ongoing civil war</td>
<td></td>
</tr>
<tr>
<td>Only geographic areas where rebels are active</td>
<td></td>
<td>Only geographic areas with at least one politically relevant ethnic group</td>
<td>Only geographic areas with at least one politically relevant ethnic group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dependent variable</td>
<td>All violence against civilians</td>
<td>Violence committed by rebel organizations</td>
<td>Violence committed by rebel organizations or governments (separate models)</td>
<td>Violence committed by ethnic rebels, infighters, or government troops (separate models)</td>
<td>Violence committed by ethnic rebels, infighters, or government troops (separate models)</td>
</tr>
</tbody>
</table>

**Model specifications**

The data were arranged as a pooled time series cross-sectional dataset with 43’000 year-hexagons (UCDP) or 344’000 month-hexagons (ACLED) for testing Kalyvas’ argument. For the main analysis, we specified models as logistic regressions with violence against civilians as the dichotomous dependent variable, standard errors clustered on the hexagon, and controls for past violence as well as for violence in neighboring hexagons that might spill over into the focal hexagon. Country fixed effects models take into account time-invariant country characteristics.
such as levels of rule of law, the security forces’ capacity to protect citizens from violence, the type of civil war, and so forth.

To check for robustness of the results to different specifications, we ran all models without fixed effects, and specified the final model with the rare events logit estimator provided by King and Zeng (2001) as well as a zero-inflated negative binomial regressions with the number of dead civilians as the dependent count variable. Finally, we also re-arranged the dataset into a cross-nested hierarchical model (with hexagons nested into countries and years). All these different model specifications produce substantially very similar results, as briefly discussed below.

Results

The dependent variable changes depending on the theory evaluated: all violence against civilians for looting models; violence committed by rebel organizations only (loose control over fighters); by rebels or governments (containing collaboration); by ethnic rebels, infighters, or government troops (ethnic targeting). We therefore need to proceed in a step-wise fashion and evaluate each theory with its own models.

In a final step, we present two models that integrate all significant variables from either the UCDP or the ACLED datasets and define any violence against civilians as the dependent variable. This will allow testing whether the results of the previous, actor-specific models hold up when regressing on all violence committed by any kind of actor. It will also allow seeing whether

---

\textsuperscript{14} Test statistics indicated that a zero-inflated model is to be preferred over others.
the ethnic targeting models can be confirmed when taking areas into account where ethnicity is
not relevant and when controlling for mechanisms identified by other theoretical approaches.

Looting

The first model in Table 2 shows no support for the idea that violence against civilians is driven
by competition for lootable natural resources (in line with Bellows and Miguel 2006] analysis of
data from Sierra Leone and Hegre et al. 2009] on Liberia). Civilians are not more likely to
become victims of violence when they live in a hexagon with many wells that produce oil, that
contains a diamond deposit, or other economically significant mineral mines. The same holds true
if we use other coding of these variables: distance to production sites, number of such sites within
a radius of 50 or 100 km, or a coding of conflict minerals (results not shown).15

15 We also don’t find any association with opium or cannabis (data from Buhaug and Lujala 2005).
Table 2: Logistic regressions on events of violence against civilians

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>All violence</th>
<th>Rebel violence</th>
<th>Low level gov. violence (ACLED)</th>
<th>Low level rebel violence (ACLED)</th>
<th>Gov. violence</th>
<th>Rebel violence</th>
<th>Infighter violence</th>
<th>All violence</th>
<th>All violence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Presence of oil producing wells (dummy)</td>
<td>-0.1758</td>
<td>-0.0697</td>
<td>-0.1702</td>
<td>-0.1683</td>
<td>-0.1625</td>
<td>-0.1469</td>
<td>-0.1468</td>
<td>-0.1471</td>
<td>-0.1471</td>
</tr>
<tr>
<td>Presence of diamond deposit (dummy)</td>
<td>0.0537</td>
<td>0.0511</td>
<td>0.0511</td>
<td>0.0511</td>
<td>0.0511</td>
<td>0.0511</td>
<td>0.0511</td>
<td>0.0511</td>
<td>0.0511</td>
</tr>
<tr>
<td>No of economically significant mining sites</td>
<td>0.0348</td>
<td>0.0348</td>
<td>0.0348</td>
<td>0.0348</td>
<td>0.0348</td>
<td>0.0348</td>
<td>0.0348</td>
<td>0.0348</td>
<td>0.0348</td>
</tr>
<tr>
<td>Strength of control over rank-and-file rebel fighters</td>
<td>0.0343</td>
<td>0.0343</td>
<td>0.0343</td>
<td>0.0343</td>
<td>0.0343</td>
<td>0.0343</td>
<td>0.0343</td>
<td>0.0343</td>
<td>0.0343</td>
</tr>
<tr>
<td>Outside financial support for rebels (dummy)</td>
<td>0.0511</td>
<td>0.0511</td>
<td>0.0511</td>
<td>0.0511</td>
<td>0.0511</td>
<td>0.0511</td>
<td>0.0511</td>
<td>0.0511</td>
<td>0.0511</td>
</tr>
<tr>
<td>Rebels funded by natural resources (dummy)</td>
<td>-1.3305***</td>
<td>-1.3305***</td>
<td>-1.3305***</td>
<td>-1.3305***</td>
<td>-1.3305***</td>
<td>-1.3305***</td>
<td>-1.3305***</td>
<td>-1.3305***</td>
<td>-1.3305***</td>
</tr>
<tr>
<td>No of times gov. gained territory</td>
<td>0.2706</td>
<td>0.2706</td>
<td>0.2706</td>
<td>0.2706</td>
<td>0.2706</td>
<td>0.2706</td>
<td>0.2706</td>
<td>0.2706</td>
<td>0.2706</td>
</tr>
<tr>
<td>No of times rebels gained territory</td>
<td>0.4668***</td>
<td>0.4668***</td>
<td>0.4668***</td>
<td>0.4668***</td>
<td>0.4668***</td>
<td>0.4668***</td>
<td>0.4668***</td>
<td>0.4668***</td>
<td>0.4668***</td>
</tr>
<tr>
<td>Distance to any territory change</td>
<td>-0.5536***</td>
<td>-0.5536***</td>
<td>-0.5536***</td>
<td>-0.5536***</td>
<td>-0.5536***</td>
<td>-0.5536***</td>
<td>-0.5536***</td>
<td>-0.5536***</td>
<td>-0.5536***</td>
</tr>
<tr>
<td>Entire hex population in rebellion</td>
<td>-0.2528</td>
<td>-0.2528</td>
<td>-0.2528</td>
<td>-0.2528</td>
<td>-0.2528</td>
<td>-0.2528</td>
<td>-0.2528</td>
<td>-0.2528</td>
<td>-0.2528</td>
</tr>
<tr>
<td>Entire hex population included in government</td>
<td>-0.8703***</td>
<td>-0.8703***</td>
<td>-0.8703***</td>
<td>-0.8703***</td>
<td>-0.8703***</td>
<td>-0.8703***</td>
<td>-0.8703***</td>
<td>-0.8703***</td>
<td>-0.8703***</td>
</tr>
<tr>
<td>Entire hex pop included and not in rebellion</td>
<td>0.5480**</td>
<td>0.5480**</td>
<td>0.5480**</td>
<td>0.5480**</td>
<td>0.5480**</td>
<td>0.5480**</td>
<td>0.5480**</td>
<td>0.5480**</td>
<td>0.5480**</td>
</tr>
<tr>
<td>Polarization btw. included and rebellious population</td>
<td>0.5566***</td>
<td>0.5566***</td>
<td>0.5566***</td>
<td>0.5566***</td>
<td>0.5566***</td>
<td>0.5566***</td>
<td>0.5566***</td>
<td>0.5566***</td>
<td>0.5566***</td>
</tr>
<tr>
<td>Km distance to nearest border between groups in conflict (loged)</td>
<td>-0.0903**</td>
<td>-0.0903**</td>
<td>-0.0903**</td>
<td>-0.0903**</td>
<td>-0.0903**</td>
<td>-0.0903**</td>
<td>-0.0903**</td>
<td>-0.0903**</td>
<td>-0.0903**</td>
</tr>
</tbody>
</table>

**Notes**
- Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.10
- A includes the following control variables: Violence last year, violence in neighboring hexagons, close to border (dummy), distance to city, distance to capital, road length, urban population, roughness of terrain
- B includes the following control variables: Distance to battle without change in territory, violence last month, violence in neighboring hexagons, urban population size, close to border (dummy), travel time to city, length of roads, sparsely vegetated surface, roughness of terrain

Observations:
- A: 40,739
- B: 24,805

Notes on numbers of observations:
- Drops hexes without at least 1 politically relevant ethnic group; drops countries and years without any violence
- Drops countries and years without conflict areas
- Drops countries with an ongoing conventional war, hexes without any violence, and hexes that changed hands between rebels and government past year
- Drops hexes without ACLED and UCDP data; drops countries and years without any violence
- Drops countries and years without any violence
Loose control over fighters

Models 2 and 3 reduce the universe of observations to those geographic areas where rebel organizations have been active in order not to misattribute violence committed by a specific rebel group to another one with different organizational characteristics and sources of funding (using UCDP GED Conflict Polygon dataset, see Croicu and Sundberg 2012). According to Model 2, the strength of hierarchical control over fighters is not associated with the propensity of a rebel organization to kill civilians.¹⁶ Financial support by outsiders is also not correlated with the chances of civilian victimization—while additional results (not shown) demonstrate that providing rebels with troop support from the outside, a sanctuary, or weapons increases their propensity to kill civilians. In other words, rebels’ military capacity to kill is more important than the source of their revenues, as maintained by the loose control theory. With regard to natural resource funding, we find the opposite of what the loose control argument predicts: Rebel organizations that profited from the sale of natural resources are less likely to kill civilians than other organizations—at least not the civilians who live close to these resources. Perhaps this is because such rebel organizations are simply not interested in establishing control over local populations because they rely less on their support.

Containing collaboration

In order to evaluate Kalyvas’ theory of violence against civilians, we now shift to the ACLED hexagon-month dataset, for the reasons explained above. The number of observations is larger

¹⁶ We obtain the same results if we include the relative military strength of rebels (in contrast to Wood 2010, who found that relatively weaker rebels kill more because they cannot provide public goods to encourage civilian support) or a variable coding whether or not they effectively control territory—two other variables that capture the military capacity of rebel organizations to kill civilians.
because the units of observations are now hexagon-months, rather than years. Model 3 indicates
that government troops do not kill civilians the more often they gained territory in a hexagon
during the last month, but the more often they lost territory to rebels, possibly killing civilians
while retreating or in areas adjacent to positions they lost (perhaps out of revenge, frustration, or
rage). Model 4 suggests that rebels also kill where they lost, rather than where they gained
territory. Note that these results remain similar if we use a two- or three-month lag or a hexagon-
year version of the dataset without any lags (results not shown).

As discussed above, the codings of independent and dependent variables used in Models 3 and 4
take the scope conditions of Kalyvas’ theory into account as well as possible: They exclude
contested territories that shifted sides more than once during the last 12 months (and where
civilians should be safer according to the theory), events with more than 5 civilian deaths (which
are unlikely to represent selective targeting), and those occurring in the midst of conventional
wars. Models 1 and 2 in Table 3 show the results when we loosen these restrictions and take all
violence in all kinds of civil wars into account. The results remain broadly similar, except that
rebel forces now also target civilians where they recently acquired territory (in line with
Kalyvas’ theory; but they still kill civilians where they lost territory as well).

What emerges from these results is that civilians are more likely to become victimized the closer
they live to actual fighting and the more they are therefore accessible to fighters. All signs of the
coefficients of the territory change variables are positive if they are significant. Even more
importantly, so is the number of battles without territory changes (not shown). In the final
models, we therefore need to control for the distance to any kind of territory change to account
for the simple effects of proximity to troops and rebels that fight each other.
Ethnic targeting

Models 5 to 7 in Table 2 evaluate the two competing ethnic targeting arguments. Each model refers to a different perpetrator-victim dyad and accordingly, we use three different dependent variables: government violence against civilians (Model 5); violence committed by rebels (Model 6); and violence by “infighters” (Model 7). We don’t find consistent support for the idea that armed groups attack their adversaries’ co-ethnics where these are most numerous, in contrast to Fjelde and Hultman (2014): Government troops don’t target hexagons entirely populated by the co-ethnics of rebels or infighters more than they do other hexagons (Model 5). Rebel fighters even significantly avoid, rather than target hexagons exclusively populated by the co-ethnics of government elites (Model 6). We arrive at the same conclusions if we use a share of the enemy’s co-ethnics as an independent variable (results not shown). Both results together allow us to conclude that neither rebels nor governments attack their enemy’s co-ethnics where they are more numerous and thus easy to target. Only in models of violence committed by infighters does the strategic logic of attacking one’s enemy’s support base appear to be at work (Model 7).

In Online Appendix 3, we offer an explanation for why our findings diverge from those of Fjelde and Hultman who use similar data sources: Regarding government violence, the differences are explained by the fact that Fjulde and Hultman do not include polarization and distance in their models. Regarding rebel violence, the findings become consistent with the ones presented here if some of their (highly collinear) controls are dropped or if we use the substantially more meaningful share of the included population as an independent variable, rather than a dummy for the presence of any included group, as they do in their models.

The ethno-territorial competition argument, on the other hand, predicts that levels of polarization between groups in rebellion and those in power should determine the likelihood of violence
against civilians. Indeed, the sign of the polarization measure is significant in the expected
direction and at the highest level in all models. The effect is also substantially important:
Increasing polarization by one standard deviation increases the probability of violence by ca. 0.45. We also find evidence for the second mechanism driving ethno-territorial competition:
Competing fighting organizations attempt to expand the continuous territory under the control of co-ethnics by focusing violence on the borders between ethnic settlement areas. The variable that measures the kilometer distance to the nearest boundary between ethnic group(s) represented in government and those in rebellion is robustly significant in all models. A one standard deviation increase in this distance decreases the probability of violence by ca. 0.2—a magnitude in the order of some of the more important control variables. Thus, in line with the ethno-territorial competition model, civilians are killed in areas that allow fighting organizations to expand the territory under control of co-ethnics where this is strategically most useful and at the same time most effective.

**Spuriousness and endogeneity**

It could be, however, that the relationship between polarization and violence is spurious because we don’t have any information on the ethnic background of victims and thus cannot exclude that an armed group attacks its own co-ethnics, rather than those of its adversary. Additional analysis (not shown here) demonstrates that “neutral” populations (groups 2-4 in Figure 1 for rebellion and groups 1-4 for infighting) are significantly less likely to be targeted by government troops and rebels (these results don’t hold up for infighters, though). We also find that neither rebels nor infighters specifically target their own co-ethnic civilians. It is therefore less likely that the association between polarization and violence is spurious.
Additional support for the idea of ethnic targeting comes from country-level studies for which information on the ethnic background of villages, which are often ethnically homogenous, is available. They show that armed organizations rarely attack their co-ethnic populations (for Darfur Olsson and Siba 2009; for Guatemala Sullivan 2012). One of the rare datasets with information on the ethnic background of individual victims comes from Northern Ireland. A full 74% of all individuals killed by Protestant paramilitaries, for example, were Catholic civilians and only 16% Protestant civilians (e.g. family members of rival paramilitaries or “traitors”).

It is also possible that polarization as well as the distance to settlement border simply pick up the difference between any mixed territory and mono-ethnic hexagons, rather than specifically a polarized ethno-demographic configuration and the continuous distance to where territories overlap. To evaluate this possibility, we coded more empirically intuitive variables indicating whether 0-10 percent of the population belongs to rebellious groups, 11-20%, 21-30%, etc. The results show that most of the government violence occurs when its co-ethnic population in a hexagon reaches between 50 and 60%. For rebel violence, a share of the co-ethnic population between 60 and 70 percent is the most violence prone. It thus seems that both government and rebels are more likely to attack civilians where their co-ethnic population is already in a slight demographic majority—quite in line with the general thrust of the ethno-territorial competition argument. Note, however, that these results look different without country fixed effects, and we therefore should perhaps not rely too much on them.

Finally, let me briefly address concerns about reverse causation. Could it be that violent cleansings produce a polarized local settlement pattern by driving out previously dominant groups from the opposite side of a conflict until they approximate half of the population—a relationship that is masked by the yearly resolution of the data? The temporal order of how
independent and dependent variables were measured makes this rather unlikely: Much of the African geo-coding of EPR relies on maps drawn by Soviet ethnographers in the 1960ies, long before the recent wave of civil wars and long before the data series on violence against civilians start in 1989. In some cases of massive population shifts, Geo-EPR offers different maps for different periods. We made sure that our ethnic settlement maps refer to periods *before* the conflict data series sets in. In other words, temporal ordering mitigates against some of the most obvious reverse causation problems.

*Final models of all violence committed by any actor in any conflict*

The preceding analysis referred to violence committed by either government, or rebels, or infighters. Correspondingly, the universe of cases varied quite dramatically. How can we bring the three different ways of behaving towards civilians into one model that uses the entire range of observations? And how can we also bring back into the picture the other variables that proved to be significantly associated with violence against civilians in previous models, in order to see whether the logic of ethno-territorial competition actually operates even when taking other mechanisms into account? And finally, does ethnic targeting show up as a significant factor of civilian victimization even if we include areas where ethnicity is not politically relevant, thus extending the sample to all countries and areas under civil war?

Models 8 and 9 in Table 2 offer such a more encompassing view. The dependent variable is *all* violence against civilians—whoever commits it. In Model 8, we include a variable that measures the distance to a territory change from the ACLED dataset because we have seen above that *any* kind of territory change (in whoever’s favor) is associated with an increased likelihood of violence. Unfortunately, combining ACLED with UCDP data would reduce the number of
observations since ACLED contains data from 1997 onward only. We therefore run a separate model with a crucial ACLED-based variable (Model 8) and another one (Model 9) with an additional UCDP based independent variable that proved to be significant in previous models: the natural resource funding of rebels.

In both Models 8 and 9, the logic of ethnic targeting is evaluated with a term that measures the degree of polarization between any group in rebellion (whether rebels or infighters) and the included population. This variable allows capturing the targeting logic of governments, rebels, and infighters alike. We now code 0 on the polarization variable for all those hexagons without any politically relevant ethnic groups, while in the ethnic targeting models (Models 5 to 7) above, those areas were dropped from the analysis. The same goes for the distance to a border between groups in conflict variable, which for the purposes of Models 8 and 9 is coded as the distance between included groups and any group in rebellion.

The results of both models are encouraging. Ethno-political polarization and distance to ethnic settlement borders are significant factors to understand violence against civilians, even if we include in our analysis countries and periods where ethnic politics is not a major driving factor, and even if we control for rebels’ capacity to kill, for natural resource funding, and for the distance to territory changes between rebels and government troops.

*Robustness checks*

For robustness purposes, we ran Model 9 as a rare event logistic regression as well as a zero-inflated negative binomial with a count of the number of killed civilians as the dependent variable. The results (see Models 3, 4, and 5 in Table 3) are substantially very similar. We conducted a second robustness check by using 50 km squares as units of observation and
recoding all variables accordingly (on the modifiable areal unit problem, see Schutte and Weidman 2011). Results (not shown here) remained substantially identical—except that in Models 6 and 8 the polarization variable does not achieve standard levels of significance since coarser units of observation loose much of the empirical information contained in the ethnic settlement maps. Third, he re-constituted the entire data-set with third-level administrative units (or counties) as units of observation in order to check whether using politically more meaningful units changes any of the main results—which is not the case (results again not shown).

Finally, we evaluated whether the results presented in Table 2 rest on a small number of observations only. For models that are restricted to hexagons in countries and years in which ethnicity is politically relevant (Models 5 to 7 in Table 2), we found that five hexagons from Mali, Uganda, Rwanda, and the DRC had either high residual or high leverage values. Removing them from the dataset did not change anything, nor did running the analysis deleting individual countries from the analysis.
### Table 3: Different model specifications

<table>
<thead>
<tr>
<th>Dependent variables</th>
<th>Diff. specifications of Models 3 and 4 in Table 2</th>
<th>Diff. specifications of Model 9 in Table 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Gov. violence (ACLED)</td>
<td>0.1705</td>
<td>0.2571***</td>
</tr>
<tr>
<td>(0.151)</td>
<td>(0.098)</td>
<td></td>
</tr>
<tr>
<td>Rebel violence (ACLED)</td>
<td>0.3344***</td>
<td>0.2169***</td>
</tr>
<tr>
<td>(0.074)</td>
<td>(0.070)</td>
<td></td>
</tr>
<tr>
<td><strong>Independent variables</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No of times government gained territory</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.1705</td>
<td>0.2571***</td>
</tr>
<tr>
<td></td>
<td>(0.151)</td>
<td>(0.098)</td>
</tr>
<tr>
<td>No of times rebels gained territory</td>
<td>0.3344***</td>
<td>0.2169***</td>
</tr>
<tr>
<td></td>
<td>(0.074)</td>
<td>(0.070)</td>
</tr>
<tr>
<td>Number of oil production sites</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.8542*</td>
<td>-1.5118***</td>
</tr>
<tr>
<td></td>
<td>(0.505)</td>
<td>(0.539)</td>
</tr>
<tr>
<td>Natural resource funding of rebels</td>
<td>-0.7738***</td>
<td>0.2495</td>
</tr>
<tr>
<td></td>
<td>(0.170)</td>
<td>(0.342)</td>
</tr>
<tr>
<td>Polarization btw. included and rebellious population</td>
<td>0.7104***</td>
<td>2.2412***</td>
</tr>
<tr>
<td></td>
<td>(0.149)</td>
<td>(0.385)</td>
</tr>
<tr>
<td>Distance to border between ethnic groups (in km, loged)</td>
<td>-0.1038***</td>
<td>-0.1891***</td>
</tr>
<tr>
<td></td>
<td>(0.022)</td>
<td>(0.054)</td>
</tr>
<tr>
<td><strong>Country fixed effects</strong></td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Control variables</strong></td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>Observations</td>
<td>314,832</td>
<td>307,476</td>
</tr>
<tr>
<td><strong>Notes on No of observations</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Drops countries and years without any violence</td>
<td></td>
</tr>
</tbody>
</table>

**Notes**

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.10

A Includes the following control variables: Violence last month, violence in neighboring hexagons, close to border (dummy), distance to city, length of roads, urban population, roughness of terrain, distance to battles without territory change

B Includes the following control variables: Violence last year, violence in neighboring hexagons, close to border (dummy), distance to city, distance to capital, road length, urban population, roughness of terrain
Conclusions

A booming strand of research on civil wars seeks to understand when civilians will become victims of intentional violence at the hands of combatants. Geo-coded datasets allow linking the occurrence of violent events to other features of the same localities and the actors operating in them. This article showed that contrary to popular perceptions and much of the policy discourse, civilians in areas rich in diamonds, oil, or conflict minerals are not more likely to become victims of violence. Furthermore, rebels funded by natural resources are less likely to kill civilians, while rebels with higher organizational capacity seem to use it to kill civilians as well—rather than to constrain the looting of their rank-and-file fighters. Finally, containing and preventing collaboration in areas not yet fully controlled by armed groups, as during the Greek civil war and in Vietnam (Kalyvas and Kocher 2009), does not seem to motivate armed groups in Africa—at least as far as we can tell given the crude measurement of territorial control that we encountered in existing data.

we do find, however, consistent support for a more political view of violence against civilians according to which it is part of a strategic struggle over control of local and national level political power—rather than of natural resources, collaborators, or fighters. In conflicts where political cleavages align with ethnic divisions, both local populations and national-level political-military organizations fight over which group is entitled and empowered to dominate the political arena. In such environments, governments and rebels have incentives to target the strategically most useful areas where the settlement areas of their affiliated populations meet and the territory continuously populated by one’s own co-ethnics can thus be expanded. They do so especially when these populations are roughly of equal size and the local balance of power can thus be tilted
in one’s favor with the least effort. This twofold dynamic of ethno-territorial competition is an important part of the puzzle to understand violence against civilians.

There are ample opportunities for future research to go beyond what has been achieved in this paper. Most importantly, causal inference is limited because we do not have information on the ethnic background of the individuals killed such that we do not know whether or not government and rebels are indeed targeting those exact populations foreseen by my theoretical model. While there are strong indications, discussed in the spuriousness and endogeneity section, that this is the case, producing a dataset with a coverage similar to ACLED or UCDP-GED that would contain more information on the individuals actually killed and those who perpetrated the violence would represent an important step forward. Obviously, assembling such a dataset amounts to a herculean task given the paucity of the corresponding information in the news reports on which event datasets have to rely.

Finally, my theory of the political dynamics of ethno-territorial competition was limited to ethnic conflicts. However, as Balcell’s work on the Spanish civil war has shown, very similar mechanisms can be found where political cleavages and rebel-civilian relationships are based on shared political ideology (see also on Afghanistan Hirose et al. 2017). It would represent a significant, but very demanding improvement if the Ethnic Power Relations dataset, including its geo-coded version, could be expanded to include other politically relevant, relatively stable political cleavages. The analysis could then go beyond deciphering the logic of ethnic targeting and reach for a more general understanding of the political dynamic underlying violence against civilians.


