Cultural Distance and Conflict-Related Sexual Violence*

Eleonora Guarnieri Ana Tur-Prats

February 2020

Abstract: This paper examines the role of ethnic-based gender norms in explaining the occurrence and intensity of sexual violence in conflict. We generate a novel dyadic dataset that contains information on the ethnic identity of the actors involved in 33 ethnic civil conflicts in Africa between 1989 and 2009 and their use of sexual violence. After exploiting ancestral economic, family, and societal arrangements, we construct and validate an ethnic-based gender inequality index. We control for a large set of fixed effects and find empirical support for two interrelated hypotheses. First, gender-unequal armed actors are more likely to be perpetrators of sexual violence. Second, we consider the perpetrator's gender norms relative to the victim's. Applying a gravity approach, we find that sexual violence is driven by a specific *clash of conceptions* on the appropriate role of men and women in society: sexual violence increases when the perpetrator is more gender-unequal than the victim. We show that (i) these patterns are specific to sexual violence and do not explain general violence within a conflict; (ii) differences in other dimensions of culture unrelated to gender do not explain conflict-related sexual violence.

Keywords: Ethnic civil conflict, sexual violence, culture, gender norms

JEL classification: D74, J16, O55, Z1.

*Author affiliations and contacts: Guarnieri (ifo Institute at the University of Munich, guarnieri@ifo.de), Tur-Prats (University of California at Merced, atur-prats@ucmerced.edu). We thank Sofia Amaral, Anke Becker, Jeff Butler, Gordon Dahl, Christian Fons-Rosen, Andreas Kotsadam, Nathan Nunn, Helmut Rainer, and seminar and conference participants at Harvard, UBC, ASREC 2019, and EconWaG Worskhop at ifo Institute, for useful comments and suggestions. Sarah Bögl provided excellent research assistance. All errors remain our own.

1. Introduction

Sexual violence in armed conflict is one of the most brutal forms of violence against women. It is a widespread crime that encompasses, among others, acts of rape, sexual slavery, and forced prostitution (International Criminal Court [2002]). At least 500,000 women were raped during the Rwandan genocide (April-July 1994), 50,000 during the Bosnian war (1992-1995), 250,000 during the Sierra Leonean civil war (1991-2002), and 400,000 in a single year of the ongoing conflict in Eastern Congo (Meger [2016]). This phenomenon comes with disastrous long-lasting consequences for victims, their families, and their communities (Ba and Bhopal [2017]). Yet, little is known about what is driving this form of violence.

Despite being widespread, armed-related sexual violence is not ubiquitous; its prevalence and intensity vary considerably both *across* and *within* conflicts (Skjelsbaek [2001]). Why do some actors systematically rape while others never do so? How is this related to the use of general (not sexual) violence? What are the fundamental drivers of conflict-related sexual violence? The aim of this paper is to shed light on these issues, and analyze the role of *ethnic-specific* gender norms in explaining the occurrence and intensity of sexual violence in armed conflicts in Africa.[1] We propose and test two hypotheses: (1) gender-unequal ethnic actors are more likely to perpetrate sexual violence, and (2) conflict-related sexual violence is explained by the cultural distance in gender norms of the combatants.

To this end, we build on the Sexual Violence in Armed Conflict (SVAC) dataset (Cohen and Nordås [2014]), which comprises all civil conflicts between 1989 and 2009 and includes an index of the intensity of sexual violence that ranges from 0 (no sexual violence) to 3 (massive and systematic sexual violence). We combine this dataset with other sources in order to include information on the ethnic identity of both actors involved and their ancestral socioeconomic characteristics. More precisely, we use the dyadic version of the SVAC dataset, GEO-SVAC (Bahgat *et al.* [2016]), to add information on the identity of both actors involved in the conflict (i.e., government/state military and rebel forces). We then assign to each actor its ethnic identity using the Ethnic Power Relations (EPR) dataset (Vogt *et al.* [2015]). Finally, we use the Murdock Ethnographic Atlas to add

[1] We restrict our analysis to Africa for two additional reasons. First, the ethnographic information is better documented and systematized for this continent. Second, the concordance table that we use to merge ethnic groups from the EPR dataset to the Murdock Ethnographic Atlas, provided by Michalopoulos and Papaioannou [2016], covers only African countries. Furthermore, Desmet *et al.* [2017] show that ethnic identity is a strong predictor of culture.

information on the ethnic characteristics of each actor. The resulting dataset has a dyadic and bidirectional structure and contains information on the intensity of sexual violence as well as on the ethnic characteristics at the actor-conflict-country-year level.[2] Our dataset covers 128 actors (106 related pairs of actors) involved in 33 ethnic civil conflicts fought in 27 different countries spanned over 20 years (1989-2009).

The empirical investigation includes several steps. First, we relate ethnic traits to anthropological notions of gender equality to construct a novel measure of gender inequality at the ethnic-group level. We find that ethnic groups characterized by ancestral arrangements conducive to more gender-unequal norms—e.g., patrilocality and dependence on male-dominated activities such as animal husbandry and pastoralism—are more likely to perpetrate sexual violence when involved in a conflict. Conversely, ethnic groups with ancestral arrangements more conducive to gender equality—e.g., matrilineality and dependence on agriculture, in which women played a prominent role—are less likely to engage in sexual violence in conflict.

Using principal component analysis (PCA), we group nine relevant ethnic traits—matrilineality, patrilocality, stem family types, dependence on gathering, hunting, agriculture, husbandry, pastoralism, and the use of the plough—into an Africa-wide ethnic gender inequality index (eGII). The first principal component explains a large fraction of the variation in the data, suggesting that economic, family, and societal characteristics across ethnic groups in Africa are intertwined. We show that these interrelated ancestral traits are powerful in capturing an ethnic group's gender norms: our eGII tallies well with contemporary measures of gender inequality such as female employment, gender attitudes, and attitudes towards intimate partner violence. The eGII is also positively associated with sexual violence in conflict, and these associations hold when including conflict fixed effects, year fixed effects, conflict-specific time trends, and victim's ethnic characteristics. In line with findings of the previous literature on gender-based violence during peacetime (Alesina *et al.* [forthcoming]; Tur-Prats [2019]), this shows that deeply-rooted gender norms also determine violence against women perpetrated in times of war.

In the second part of our empirical analysis, we test whether sexual violence is explained by the cultural distance in gender norms between the combatants. We exploit the dyadic

[2]For example, for the conflict that in 1994 confronted the Government of Chad against the *Comité de Sursaut National pour la Paix et la Démocratie* (CSNPD), we observe the prevalence and intensity of sexual violence that the Government of Chad exerted against the CSNPD (intensity 2) and the sexual violence that the CSNPD exerted against the Government of Chad (zero). At the same time, we also observe the ethnicity of both actors: the rulers of the Government of Chad at that time came from the Zaghawa and Bideyat ethnic groups, while the CSNPD rebel forces were formed by Sara soldiers. Table 1 illustrates this.

structure of the data, and run a specification in the spirit of a gravity equation, similarly to recent literature adjusting the canonical trade models to other contexts (Becker *et al.* [2018]; Grosjean [2011]; Serafinelli and Tabellini [2017]; Spolaore and Wacziarg [2009]). By regressing the intensity of sexual violence on the absolute distance in gender norms between perpetrator and victim, we find that the larger the cultural distance between the ethnic belligerents, the higher the intensity of sexual violence.

When examining this association further, we uncover that this effect is driven by a specific *cultural clash*, one where the perpetrator holds more gender-unequal norms than the victim. These results remain significant when including conflict fixed effects, year fixed effects, and conflict-specific time trends. Our results also hold when we separately include perpetrator and victim fixed effects, i.e., when we isolate the effect of cultural distance from the combatant's own characteristics, including gender norms, aggressive nature, and female vulnerability, among others. Furthermore, our results are robust to the inclusion of additional sets of fixed effects (conflict-year, country, among others), to alternative versions of the eGII, and to abstracting from the temporal variation in the data.

We then discuss different potential explanations for our results and are able to rule out two of them: (1) that perpetrators strategically use sexual violence to destroy women in the opponent's society since they represent a valuable asset, and (2) that women in some ethnic groups might be more vulnerable and thus more likely to be victimized. We instead offer a new explanation based on the identity model by Akerlof and Kranton (2000) which has the advantage of explaining why we only find positive effects of cultural distance when perpetrators are more gender-unequal than their victims. When confronting a more gender-equal society, perpetrators might perceive the relatively better position of women as a threat to their own ideals, and resort to sexual violence to restore the loss of utility experienced by this encounter.

We show that these patterns are specific to gender-based violence. Cultural distance in gender norms can be correlated with other dissimilarities between ethnic groups, which in turn could trigger violence. If this were the case, then sexual violence could be the byproduct of general violence, and the cultural clash that we are examining would not be necessarily linked to a gender-based form of violence. To test this, we conduct a placebo exercise in which we re-run our analysis replacing sexual violence with a measure of general violence: the number of deaths inflicted by the perpetrator on the victim. Because this measure is also bidirectional, we can replicate our cultural distance analysis using this alternative dependent variable. Our results show that cultural distance in gender norms

does not explain general violence within a conflict.

Finally, we analyze other dimensions of cultural differences not directly related to gender. To this end, we exploit two widely-used measures of cultural distances between populations: linguistic and religious distance. Following Fearon [2003], we first construct a measure of linguistic distance for each related pair of actors and test whether a clash of more general cultural conceptions among populations can also explain sexual violence in conflict. Even though distance in gender norms is positively correlated with linguistic distance, we do not find that linguistic distance is associated with sexual violence. This finding is robust to alternative specifications in which for instance we isolate the component of cultural distance that is unexplained by differences on gender norms.[3] Similarly, we do not find evidence that religious differences between combatants are powerful in explaining the use of sexual violence: controlling for religious distance in our preferred specification leaves the main coefficients unchanged and, if at all, religious distance between perpetrator and victim is negatively associated with sexual violence. Taken together, these results suggest that conflict-related sexual violence is not driven by general cultural differences, and that its understanding requires a gender-based explanation.

This paper is related to several strands of literature. First, we contribute to the interdisciplinary literature on sexual violence in armed conflict—summarized in the next Section—by advancing and empirically testing a new hypothesis, namely that the prevalence and the intensity of war-related sexual violence is explained by a *clash of conceptions* on what is the appropriate role of men and women in society. Second, this study adds to the literature on gender norms and gender inequality by proposing and validating a new gender inequality index based on ethnic traits, and enriches the growing literature on the long-run cultural determinants of violence against women (Alesina *et al.* [forthcoming], Tur-Prats [2019]) by empirically investigating a so-far overlooked form of gender-based violence. Third, our paper is related to the vast literature on how ancestral conditions, by persistently shaping cultural norms, can influence contemporary outcomes and behaviors (Alesina *et al.* [2013]; Becker [2018]; Guiso *et al.* [2016]; Voigtländer and Voth [2012]).

Finally, we contribute to work on ethnic conflict in developing countries (see Blattman and Miguel [2010] for a summary), and we do so in two ways. First, we construct a novel dyadic dataset, which includes information on the ethnic identity and ancestral characteristics of *all* actors involved in an ethnic civil conflict. Second, while the literature has mostly focused on understanding whether and how diversity triggers war (see next

[3] We first regress linguistic distance on distance in gender norms, and use the residuals of this regression as our main explanatory variable.

Section for a summary), we hypothesize and show that cultural distance between actors can also explain how violence unfolds once conflict takes place.

The remainder of the paper is structured as follows. Section 2 reviews the related literature and advances our hypotheses. Section 3 describes the data and the procedure used to merge the various sources, and Section 4 presents descriptive statistics. In Section 5 we discuss the empirical strategy for investigating the impact of the perpetrator's gender norms on the use of sexual violence, and in Section 6 we report the results of this exercise. Section 7 describes the ethnic-based gender inequality index and tests the cultural clash hypothesis, and finally Section 8 concludes.

2. Existing Literature and Hypotheses

The first hypothesis we test is that gender-unequal ethnic actors are more likely to perpetrate sexual violence during a conflict. The idea that gender norms in a society and the prevalence of gender-based violence are linked has been advocated by scholars across various disciplines. Two contributions in the economics literature have empirically shown that deeply entrenched norms about the role of men and women in society are associated to intimate partner violence (IPV). In Africa, women belonging to ethnic groups where, in ancestral times, women had more marginalized roles in the economy and society relative to men are today more likely to be IPV victims (Alesina *et al.* [forthcoming]). In Spain, Tur-Prats [2019] finds that historical family structures—stem versus nuclear—influenced women's participation in non-domestic work and persistently shaped gender roles. Areas where stem families were predominant in the past are characterized by more progressive gender norms, and women residing in these regions are today less likely to report IPV.

Among scholars investigating the determinants of gender-based violence in the context of armed conflict, some favored a socio-cultural rationale behind soldiers' use of sexual violence. Through the analysis of previous literature, Skjelsbaek's [2001] noticed a consensus in considering sexual violence a weapon of war, i.e., part of a pre-meditated strategy, rather than the manifestation of a latent biological need triggered by a state of war.[4] In Skjelsbaek's [2001] conceptualization, perpetrators use sexual violence strategically to empower (i.e., masculinize) their own identity and to victimize (i.e., feminize) the opponent's. According to this view, any attempt to analyze sexual violence in conflict without

[4]Inspired by early anthropological work by Symons [1979], a much-discussed contribution by Thornhill and Palmer [2001] describes rape through the lenses of evolutionary biology. According to the authors, rape is a biologically-determined behavior: it is either the direct result of an evolutionary adaptation to increase men's reproductive success, or a byproduct of other adaptive traits, such as aggressiveness.

considering gender relations is incomplete. Meger [2016] also lists context-specific gender norms—in turn shaped by political, economic, and social structures—as one of the factors underlying the occurrence of conflict-related sexual violence. These views, in turn, are in line with anthropological research on sexual coercion. For example, Sanday [1981b] suggests that rape is an expression of cultural forces operating at the societal level. Through the analysis of a cross-cultural sample of tribal societies, she shows that rape-prone societies are characterized, among other things, by a higher degree of male dominance, compared to non rape-prone ones.

The first empirical analysis of the determinants of sexual violence in armed conflict was conducted by Cohen [2013], who finds support for the so-called *combatant socialization theory*, according to which soldiers recruited by force use rape as a method to socialize and generate cohesion. Contrary to the conjectures of the aforementioned literature, Cohen [2013] does not find a relationship between gender inequality and sexual violence in conflict at the country level. Albeit positive, the correlation between rape and fertility—used as a proxy for gender inequality—is not statistically significant. We test the gender inequality hypothesis by moving from the *country level* to the *conflict-actor's level*, and by measuring gender inequality through ethnic actors' deeply-rooted cultural norms, in turn shaped by their ancestral economic, societal, and family arrangements.

After finding empirical support for the first hypothesis, we take a step forward and investigate whether sexual violence is only explained by the perpetrator's gender norms, or whether the interaction between perpetrator's and victim's cultures also plays a role. We advance and test the following additional hypothesis: sexual violence increases with cultural distance in gender norms between two opposing ethnic belligerents. To the best of our knowledge, there is no existing theory or empirical evidence on how cultural distance in gender norms between ethnic groups might influence their decision to perpetrate sexual violence during a conflict. However, anthropological work by Taylor [1999] emphasizes the gender component of the violence that unfolded during the Rwandan genocide, which materialized in systematic acts of sexual violence perpetrated by Hutu men against Tutsi women. According to Taylor [1999], this was the result of Hutu men disagreeing with Tutsi women's prominent role in society:

> Hutu extremists aimed at reclaiming the lost ground of patriarchy and re-asserting a male dominance that had probably never existed in Rwanda's actual history. [...] The Rwandan genocide was not simply a battle for political supremacy between groups of men, it was also about re-configuring gender.

> [...] Gender relations were falling into a state of decadence and disorder as more [Tutsi] women attained positions of prominence in economic and public life.

Furthermore, our hypothesis is grounded in previous literature that analyzes how cultural distance between two entities can trigger a range of violent manifestations, from discrimination (see Becker's [1957] seminal work) to conflict. At the broad macro-cultural level, Huntington's [2000] 'clash of civilizations' thesis states that cultural and religious differences are the main determinants of conflict in the post World War II era. Focusing on interstate wars, Bremer [2000] argues that more ethnically distant societies will be more likely to fight against each other. Closer to our study, Caselli and Coleman's [2013] model of ethnic conflict predicts that ethnic groups are more likely to clash when the differences between them are more pronounced. Based on these theoretical insights, we hypothesize that cultural distance between ethnic groups may not only trigger conflict, but also influence in what ways violence unfolds once conflict takes place.[5]

Although our empirical analysis will not allow us to tease out precise mechanisms, the association between cultural distance and sexual violence can be rationalized as follows: deep disagreements between groups on the appropriate role of men and women in society may lead actors to strategically resort to a form of violence that directly targets and disrupts the object of this cultural disagreement, i.e. the opponent's views about gender relations. Since it is reasonable to expect sexual violence to disrupt cultural notions of gender equality, we conjecture (and find) that sexual violence is driven by a specific *cultural clash* between combatants, one where the perpetrator holds more gender-unequal norms than the victim.

The following section describes in detail the data sources and the unit of analysis we adopt to test these two hypotheses.

3. Data

We construct a novel dataset, which combines a variety of sources on ethnic conflict, the actors involved, their use of war-related sexual violence, and their ancestral ethnic characteristics. This Section provides an overview of the main data sources used for the

[5]Empirical findings on the relationship between cultural distance and the onset of conflict are mixed, and suggest that the direction of this association may depend on the nature of conflicts. When looking at intrastate conflicts, Arbatli *et al.* [2013] find that genetically diverse countries are more likely to engage in civil war. Conversely, Spolaore and Wacziarg [2016] show that genetic distance between any two countries is associated with less international conflict with each other.

analysis. More details on the data sources as well as on the procedure we adopted to construct the dataset can be found at the Section A-1 in the Appendix.

3.1. Sexual Violence in Armed Conflict

The source of our dependent variable is the Sexual Violence in Armed Conflict (SVAC) Dataset (Cohen and Nordås [2014]), which contains information on sexual violence used in civil conflicts fought between 1989 and 2009. We focus on *ethnic* civil conflicts, which are defined as "armed conflicts between the government of a state and one or more internal opposition group(s) that cause at least 25 battle-related deaths within a year and in which armed groups (i) explicitly pursue ethno-nationalist aims, motivations, and interests; and (ii) recruit fighters and forge alliances on the basis of ethnic affiliation" (Gleditsch *et al.* [2002]). We exploit the *dyadic* version of the SVAC dataset, called GEO-SVAC (Bahgat *et al.* [2016]), which includes both the identity of the perpetrator of sexual violence and the identity of the other actor involved in the conflict. Consistently with the definition of civil conflict, one of these two actors is always the government of a state, and the opponent is always a rebel group.

Adhering to the International Criminal Court's rationale, SVAC defines war-related sexual violence as including acts of rape, sexual slavery, forced prostitution, forced pregnancy, forced sterilization, and forced abortion (International Criminal Court [2000]). In addition, following Wood [2009], sexual mutilation and sexual torture are also included. SVAC draws upon annual reports from three sources (Amnesty International, Human Rights Watch, and the US State Department) to construct a measure of prevalence of sexual violence at the conflict-actor-year level. The resulting variable is an index ranging between 0-3 that reflects the magnitude of the phenomenon. More specifically, it takes the value 3 if, in a given year of conflict, an actor perpetrated acts of massive, innumerable, or systematic sexual violence and if reported incidents or victims of sexual violence exceeded 1,000; 2 if sexual violence was described as widespread and common, and reports of victims or incidents ranged between 25 and 999; 1 if reported victims and incidents were below 25 and the occurrence of sexual violence was only isolated; 0 if no sexual violence was mentioned in a given year in relation to a specific conflict.[6]

[6]For further details on the methodology of data collection and coding refer to Cohen and Nordås [2014].

3.2. Conflict Actors' Ethnic Identity and Ancestral Characteristics

To assign an ethnic identity to each conflict actor—rebel groups and governments—we exploit the rich information provided by the Ethnic Power Relations (EPR) Dataset Family (Vogt *et al.* [2015]). EPR defines an ethnic group as "an identity group that defines itself or is defined by others along linguistic, religious, or racial characteristics".

The EPR dataset family contains information, *inter alia*, on ethnic groups' involvement in civil war as part of a rebel organization and on ethnic groups' access to executive government power. We are therefore able to link each rebel force and each government to one or multiple EPR ethnic groups, depending on whether rebels and governments are the result of ethnic alliances.[7] Section A-14. in the Appendix illustrates this merging procedure with a concrete example of a conflict event in Liberia.

We add information on ethnic groups' ancestral characteristics using the Ethnographic Atlas (EA), coded by Murdock [1967] and updated by Nunn and Wantchekon [2011]. The EA is arguably the most compelling source of ethnographic information for 1,265 societies around the world, collected at the end of the 19th century. For Africa, the EA provides detailed information on groups' socio-economic conditions, settlement patterns, and family arrangements prior to European contact. We will describe these variables in detail in Section 6.

We add the information provided by the EA to the dataset on conflict through the concordance table provided by Michalopoulos and Papaioannou [2016], which links 196 EPR groups to 593 ethnicities in the EA using a variety of sources. In some cases, this matching procedure results in a one-to-one mapping between EPR groups and EA groups. For example, the ethnic group of the rebel force FLEC-FAC in Angola, the *Cabindan Mayombe*, is matched with the *Yombe* group in the EA. However, in other cases, a conflict actor is associated to multiple EA groups either because (i) an actor is represented by multiple EPR groups, (ii) an EPR group corresponds to multiple groups in the EA, or (iii) both.

An example of the latter case is the following: the RFDG rebel group in Guinea is composed of members belonging to the EPR groups called *Malinke* and *Peul*. In turn, the Michalopoulos and Papaioannou's [2016] correspondence table matches *Malinke* to four EA groups (*Yalunka, Konyanke, Malinke,* and *Koranko*), and the *Peul* to three EA groups

[7] We assume that state and rebel military forces mirror the ethnic composition of governments and rebel groups, respectively. In Section 7.5 we conduct a robustness test in which we assume that state forces mirror the ethnic composition of the entire country (weighted by the size of each group's settlement area) and show that our results are robust to this alternative definition.

(*Foutadjalon, Sokoto, Liptako*). In these instances, we weight the ethnic characteristics of each EPR group by the size of the EA groups to which it corresponds. In the just-mentioned example, *Peul*'s dependence on pastoralism will be a weighted average between *Foutadjalon's, Sokoto's*, and *Liptako's* dependence on pastoralism, based on the three ethnic groups' size, in turn proxied by the land area covered by their settlements. We will provide estimates using both the weighted and the un-weighted version of the ethnic characteristics, and show that our results are generally insensitive to this procedure.

The final sample used for the analysis includes 33 ethnic civil conflicts fought between 1989 and 2009 in 27 African countries, involving 128 different actors (106 related pairs of actors).

4. Descriptive Statistics

Sexual violence was present, in some level of intensity, in 82% of the conflicts included in our sample. 21% of all ethnic civil conflicts in Africa between 1989 and 2009 experienced at least one episode of sexual violence at the highest intensity, i.e., involving at least 1,000 victims. State forces perpetrate sexual violence more frequently than rebel groups. However, when perpetrated by rebel groups, the intensity of sexual violence is on average higher.

Figure 2 reports the spatial variation of sexual violence at the ethnicity level, conditional on the ethnic group being involved in a conflict. The striking picture that emerges is that there is considerable within-country variation in whether or not ethnic groups (and therefore, armed actors) engage in sexual violence. An interesting example is that of Algeria. In the long civil war between the government and various rebel armed forces, which began in 1991, rebel groups never made use of sexual violence, while the government constantly engaged in it throughout the war. Similarly, in the Oromo conflict in Ethiopia, the Oromo Liberation Front (OLF) did not engage in sexual violence, while the government armed forces did so repeatedly. In other cases, such as the one of Sudan, the vast majority of ethnic groups involved in conflict perpetrated sexual violence, but there was a quite large variation in the intensity of it. Finally, as in the case of Nigeria, the use of sexual violence was widespread across groups, and its incidence homogeneous.

5. Sexual Violence and Gender Norms: Empirical Strategy

In the first step of our empirical analysis, we test the relationship between the use of sexual violence in conflict and the perpetrators' ancestral socio-economic characteristics

as proxies for their gender norms. We do so by estimating the following:

$$SVAC_{ict} = \alpha + \beta Ethn_i + \eta_c + \phi_t + \omega_c t + \epsilon_i \quad (1)$$

where the dependent variable, $SVAC_{ict}$, denotes the intensity of sexual violence perpetrated by actor i during conflict c in year t. $Ethn_i$ captures the ethnic ancestral characteristics of the perpetrator. η_c denotes conflict fixed effects, which account for time-invariant characteristics at the conflict level (e.g., conflict motives, external support, overall conflict cruelty, type of warfare, available technology, military tactics). Year fixed effects (ϕ_t) allow to control for time-specific shocks in the whole continent (e.g., the recognition of sexual violence in conflict as a crime, international policies or protocols that might affect data collection and categorization). A conflict-specific year trend (ϕ_t) accounts for time-varying factors at the conflict level (e.g., escalation of violence). Standard errors are clustered at the perpetrator level.

The just-described estimating equation abstracts from the victim's characteristics, focusing exclusively on the perpetrator. In order to isolate the perpetrators' traits from the victim's, we also estimate the following:

$$SVAC_{ijct} = \alpha + \beta Ethn_i + \beta Ethn_j + \eta_c + \phi_t + \omega_c t + \epsilon_i \quad (2)$$

where $SVAC_{ict}$ denotes the intensity of sexual violence perpetrated by actor i against actor j on conflict c and year t, and $Ethn_j$ controls for the victim's ethnic characteristics.

6. Results

An established literature spanning different disciplines has demonstrated that ancestral economic and societal arrangements have persistently shaped gender relations. In this section, we test whether ethnic groups with ancestral traits that have been associated with gender inequality are more likely to be perpetrators of sexual violence in armed conflict. Our choice of ethnic characteristics that we include in the analysis is informed by what the literature has highlighted as relevant determinants of gender norms in a society. We focus on characteristics reflecting descent practices, residence patterns, family arrangements, subsistence activities, and exposure to the slave trade. Ethnic characteristics on which the literature is inconclusive are not included in our analysis.[8]

[8]For example, it is unclear whether the practice of brideprice increases or decreases gender inequality. On the one hand, brideprice is a recognition of women's value, and it is more typical in societies where women have an important role in agricultural production (Boserup [1970]). On the other hand, the obligation of women to pay back the brideprice in case of divorce may decrease their bargaining power.

6.1. Descent, Residence and Family Arrangements

We first explore the role of descent, residence patterns, and family arrangements in explaining ethnic groups' use of sexual violence in conflict. Anthropologists have argued that societies where these arrangements are centered around women tend to be characterized by higher gender equality (Martin and Voorhies [1975], Sanday [1981a]).

In matrilineal societies, inheritance is traced through female family members. Therefore, women are key for determining descent and have constant support from their kin network (Schneider and Gough [1961]). Lowes [2017] has shown that women belonging to matrilineal ethnic groups in the Democratic Republic of Congo detain higher bargaining power within the household compared to their counterparts in patrilineal societies, and that they are less likely to be victims of intimate partner violence. As shown in Gneezy *et al.* [2009], matrilineal women are also more likely to display behavioral traits that are usually typical of men—such as willingness to compete—ones that have often been advocated as factors explaining economic disparities between men and women. Taken together, this evidence supports the notion that matrilineal societies, when compared to patrilineal ones, are characterized by more equitable gender norms. Estimates of β in Table 2 show that matrilineal ethnic groups are less likely to engage in sexual violence in armed conflict as opposed to patrilineal or bilateral groups.[9] The intensity of sexual violence, measured on a 0-3 scale, is on average between 0.66 and 0.85 lower when perpetrated by a matrilineal armed actor. Equivalently, one standard deviation increase in matrilineality among ethnic groups forming a conflict actor decreases the intensity of perpetrated sexual violence by 0.22-0.28 standard deviations.[10]

Lineage systems and kinship structures in a society are inextricably linked to residence patterns. Patrilineal societies are also likely to be patrilocal, a system of postmarital residence where the newly formed couple moves near the husband's kin group (Murdock [1967]). In these societies, women may be less protected by their own family, and hus-

The association between polygyny and gender equality is also ambiguous. On the one hand, women's status in polygynous unions may be lower, in particular for younger wives due to early marriage and large age gaps with the husband. Alesina *et al.* [forthcoming] show that women in polygynous unions are more likely to suffer from intimate partner violence across Africa, but, at the same time, women and men in societies that traditionally practice polygamy are less likely to justify intimate partner violence. Furthermore, as highlighted in Boserup's [1970], polygyny is more typical in societies where women constitute an important economic asset.

[9]Societies practicing bilateral descent attach equivalent importance to the female and male clans.

[10]A point worth noting: variables that were originally dichotomous in the Ethnographic Atlas, such as "matrilineal" or "patrilocal", are continuous in our dataset, since one conflict actor may be formed by multiple ethnic groups that may not share all ethnic characteristics. Therefore, the interpretation of the coefficient in terms of standard deviations is meaningful.

bands may more easily exercise their authority over women. Scholars in anthropology have hypothesized that patrilocality is a direct consequence of women's low economic participation (Korotayev [2003]). Not surprisingly, conflict actors from traditionally patrilocal societies perpetrate violence at higher intensity than their matrilocal or neolocal counterparts.[11] As shown in Table 2, the magnitude of the coefficient is comparable to the one on matrilineality, but reversed in sign. One standard deviation increase in the degree of patrilocality in a conflict actor increases sexual violence by 0.21 standard deviations.

Finally, we examine the role of stem families, a family type that a recent contribution by Tur-Prats [2019] has linked to higher gender equality. The co-residence of the wife with the mother-in-law frees up women from the burden of domestic work, and allows them to exercise a productive role in the economy and participate in family subsistence. This ethnic trait is associated with a lower intensity of sexual violence in conflict. However, the magnitude of the coefficient is smaller than in the case of matrilineality and patrilocality, and significant at the 10 percent level only in some specifications.

6.2. Subsistence Activities

Next, we explore the role of subsistence activities. In ancestral societies, the relative participation of women and men in economic activities has persistently shaped gender relations (Friedl 1975, Sanday 1972). According to Friedl [1978], in hunter-gatherer societies men exert control over animal protein, a scarce and hard to acquire resource. Since hunting activities require a certain amount of physical strength, they are predominantly a men's task. As a result, these societies tend to be characterized by high degrees of male dominance. Across Africa, women in societies that traditionally relied on hunting activities are today more likely to experience intimate partner violence (Alesina *et al.* [forthcoming]). Similarly, our results in Table 3 indicate that armed actors whose ethnic groups relied on gathering or hunting are more likely to perpetrate sexual violence in armed conflict. However, albeit positive in all specifications, these coefficients are imprecisely estimated.

Conversely, ancestral dependence on agriculture is negatively associated with the use of sexual violence in conflict, as shown in the last panel of Table 3. An ethnic group's full reliance on agriculture for subsistence, compared to no dependence at all, decreases the intensity of sexual violence by 1.34-1.47 when measured on a 0-3 scale. Differently stated, one standard deviation increase in an ethnic group's dependence on agriculture

[11] In societies practicing neolocal residence the couple moves to a new physical place, i.e., neither close to the bride nor to the groom's family or kin group.

decreases the intensity of sexual violence by 0.26 standard deviations. In the African context, where agriculture was characterized by shifting cultivation, female participation in agricultural activities was traditionally high, as emphasized by Boserup [1970] and confirmed in the ethnographic data by Murdock [1967]. In contrast, in other regions of the world where plough agriculture was more common, the traditional division of labor was reversed, with men taking up the majority of agricultural work and women remaining confined to the domestic sphere. This agricultural system based on plough agriculture—and the consequent division of labor—contributed to the evolution of gender-unequal norms, as empirically demonstrated in Alesina *et al.* [2013]. In our sample, the coefficient on plough use is positive in most specifications displayed in Table 4, but not statistically different from zero.

Finally, conflict actors traditionally relying on animal husbandry or pastoralism for subsistence perpetrate sexual violence at higher intensity. As shown in Table 4, one standard deviation increase in dependence on animal husbandry or pastoralism increases the intensity of sexual violence by 0.36 standard deviations. Pastoralism is a specific type of animal husbandry based on herd animals that require natural pasture,and entailed frequent and extended periods of male absence from the community, resulting in higher paternity uncertainty. Becker [2018] shows that these byproducts of pastoralism—male absence and paternity uncertainty—incentivized the adoption of measures to control women's sexuality and mobility. Women in societies where pastoralism was historically an important source of subsistence are today more likely to be infibulated, to be restricted in their mobility, and to hold more gender-unequal attitudes. Consistently with Becker's [2018] findings, our results show that pastoralism is also an important determinant of the use of sexual violence in armed conflict.

6.3. The Slave Trade

Finally, we test whether an ethnic group's exposure to the transatlantic and Indian Ocean slave trade is a factor explaining the use of sexual violence in conflict. According to Teso [2018], exposure to a demographic shock such as the transatlantic slave trade, where slaves exported were predominantly men, contributed to the evolution of more gender-equitable norms. In heavily raided ethnic groups, women started taking up typically men's tasks, and this resulted in a shift of the traditional gender division of labor. This shock had persistent effects in the long run: today, women whose ancestors were exposed to the transatlantic slave trade are more likely to be in the labor force, and to have lower fertility and higher decision-making power within the household.

We rely on information on the number of slave shipments provided by Nunn and Wantchekon [2011] to construct a measure of exposure to the transatlantic and the Indian Ocean slave trade. As shown in Table 5, ethnic groups exposed to the transatlantic slave trade are less likely to engage in sexual violence in conflict. Moving from no slave trade exposure to the highest level of exposure in the sample decreases the intensity of sexual violence by 0.86-1.56. Equivalently, one standard deviation increase in exposure to the slave trade reduces sexual violence by 0.09-0.16 standard deviations.

Ethnic groups exposed to the Indian Ocean slave trade, conversely, use sexual violence in conflict at a higher intensity. The Indian Ocean slave trade did not distort the sex ratio as the transatlantic trade did, because it did not preferentially export men. The coefficient, however, loses significance once we control for the victim's slave trade exposure.

6.4. Robustness Tests

The just-discussed associations between ethnic traits and the use of sexual violence in conflict are robust to various alternative specifications. Results are reported in Tables B-1 to B-4 in the Appendix. First, to fully account for the victim's characteristics, we include victim fixed effects in the main specification. Columns (1) and (2) in Tables B-1 to B-4 display coefficients of this—more demanding—specification, for the weighted and unweighted version of each ethnic characteristic, respectively. The significance of the estimates tends to fall, although the majority of coefficients maintain the expected sign.

Next, we run our specification replacing conflict fixed effects with country fixed effects. This specification is less conservative than our preferred one, since one country may experience multiple conflicts.[12] However, in few instances, one conflict may span across multiple countries.[13] As columns (3) and (4) in Appendix Tables B-1 to B-4 show, the coefficients are insensitive to the choice of fixed effects, and maintain the same magnitude and significance as in the main specification.

Finally, in columns (5) and (6) we abstract from the temporal variation present in our data. Since the independent variables—i.e., ethnic ancestral characteristics—are time-invariant, we collapse the data and have as unit of observation a dyad (perpetrator-victim pair) in a specific conflict and country. The outcome variable is the average sexual violence

[12]Between 1989 and 2009, for example, Niger experienced what UCDP-GED defines as three different conflicts, i.e., the first, second, and third Tuareg rebellions, respectively, fought by five different rebel groups against the government.

[13]For example, the conflict between the government of the Central African Republic and the Forces of Francoise Bozize took place both in the Central African Republic and in Chad. Similarly, some events in the conflict between the government of Ethiopia and the Oromo Liberation Front took place in Kenya.

intensity perpetrated by each actor in all years of a specific conflict. Reassuringly, results are similar to those obtained with the specification that includes the temporal variation.

7. Testing the Cultural Clash

As shown in the previous section, perpetrators of sexual violence in conflict are characterized by ethnic traits that have been associated with notions of gender inequality. In the second step of the empirical analysis, we test our second hypothesis: does the cultural distance in gender norms of the combatants help explain the emergence of sexual violence in ethnic conflicts? Or, differently stated, are armed actors more likely to use sexual violence when confronted with opponents that hold different cultural norms on the appropriate role of women and men in society?

So far, we have separately analyzed the role played by each ethnic trait in explaining sexual violence. To obtain a measure of cultural distance between conflict actors, we construct a one-dimensional indicator capturing gender norms in a specific ethnic group. We group all the ethnic traits analyzed in the previous section into an ethnic gender inequality index (eGII), which increases in the prevalence of ethnic traits associated with gender inequality. The next section describes the procedure we adopt to construct the index.

7.1. Ethnic Gender Inequality Index

Among those ethnic characteristics that have been associated with gender (in)equality in the literature and based on our results, we choose nine traits included in the Ethnographic Atlas and construct an Africa-wide index using principal component analysis (PCA). [14] Three of these nine traits—matrilineality, stem family, and dependence on agriculture—are reconcilable with notions of gender equality and negatively associated with sexual violence in armed conflict. Therefore, we expect these single traits to be *negatively* correlated with the eGII. The remaining six ethnic characteristics—the use of the plough, patrilocality, dependence on gathering and hunting, dependence on pastoralism, and animal husbandry—have been associated with gender inequality and with higher use of sexual violence in conflict. In turn, we expect these single traits to be *positively* correlated with the eGII.

The first principal component alone explains 32% of the common variance of the nine

[14] We do not include the slave trade in the analysis, due to the fact that it was a geographically constrained historical shock.

ethnic characteristics across Africa. Table 6 presents the loadings of each of the nine traits included, i.e., their correlation with the first principal component. As can be seen, the sign of the loadings is as expected for the majority of ethnic traits. Matrilineality and dependence on agriculture are negatively correlated with the first component, while patrilocality, dependence on pastoralism, the use of the plough, and dependence on animal husbandry are positively correlated. Instead, the correlation between stem families and the first component is particularly small, and with an unexpected sign. Similarly, dependence on hunting and gathering are negatively correlated with the fist principal component, despite their positive (but not significant) association with sexual violence in conflict and despite what is argued in the anthropological literature.

Due to the presence of these "ambiguous" ethnic traits, we also provide an alternative version of the eGII, i.e., a restricted one based exclusively on ethnic characteristics that are unambiguously linked to gender (in)equality and, in addition, are robustly associated with sexual violence in conflict.[15] Table B-6 in the Appendix reports the corresponding loadings for these five ethnic traits, which hold the expected sign. We show that our results are robust to the use of this alternative version of the eGII.

We normalize the predicted score of the PCA to range between 0 and 1, with 0 denoting highest gender equality and 1 denoting highest gender inequality. Figures 3 and 4 report the distribution of the eGII across Africa and in our sample, respectively. Ethnic groups in our sample, i.e. those that were involved in at least one conflict between 1989 and 2009, tend to be characterized by more gender unequal norms on average, compared to the average of the continent as a whole.

Figure 5 reports the geographical distribution of the eGII across Africa, displaying the Murdock ethnic map and the corresponding eGII for each group. The highest levels of gender inequality are prevalent in ethnic groups located in North and East Africa, while the lowest are concentrated in Central Africa and in some parts of West Africa. The distribution of the eGII varies considerably across regions, but also within countries. One of the most extreme cases is Tanzania, where ethnic groups span from the lowest bin of the eGII (0-0.25) to the highest bin (0.75-1). Figures B-2 and B-1 in the Appendix display the distribution of the restricted version of the index. The distribution across Africa is very similar to the one of the main eGII.

As shown in Figure 5, our eGII is correlated with proxies for gender (in)equality today. Countries with the lowest rates of female labor force participation are also those in which

[15]We exclude the dependence on gathering and hunting, stem family, and the use of the plough, the latter one due to lack of statistical significance of the plough coefficients in our analysis.

ethnic groups are characterized by a high eGII. This correlation is also found at the ethnicity level. Figures 6 to 13 report the correlation between the eGII and various indicators of gender (in)equality using micro-data from the Demographic and Health Survey (DHS) and the Afrobarometer survey. We match individuals to the Murdock Atlas via latitude and longitude.[16] Figure 6 shows that ethnic groups in sub-Saharan Africa where female employment is lower display higher values of the eGII. The eGII is also correlated with attitudes towards intimate partner violence (Figure 7): in ethnicities where the eGII is higher, respondents are more likely to justify violence, and report at least one reason for which the husband is justified to beat his wife.

Data from the Afrobarometer survey shows a correlation with gender attitudes. Figure 8 shows that ethnic groups with a higher eGII are less likely to agree with the statement that "women should have equal rights and receive the same treatment as men do". Along the same lines, ethnic groups with a high eGII are more likely to hold unequal gender attitudes, elicited through the agreement with the statements that "men make better political leaders than women, and should be elected rather than women" (Figure 9) or "if funds for schooling are limited, a boy should always receive an education in school before a girl" (Figure 10). Respondents belonging to more gender-unequal ethnic groups are also more likely to report that women are always or often treated unequally by leaders, the police, or the employer (see Figures 11-13). These correlations tend to hold also for the restricted version of the eGII, as shown in Figures B-3 and B-4 in the Appendix. Taken together, this suggests that the eGII—which embeds information on ancestral arrangements that may no longer be in place today—performs fairly well in capturing contemporary notions of gender inequality.

Finally, Figure 14 compares the distribution of the eGII with the distribution of sexual violence in armed conflict. We re-run the Equation 1 using the eGII instead and, not surprisingly, find that the eGII is positively and significantly associated with sexual violence (Table 7). In our preferred specification in column (2), one standard deviation increase in the eGII increases sexual violence by 0.45 standard deviations (0.36 when controlling for the victim's eGII). These results hold when running the same robustness tests proposed for individual ethnic traits, and when using the restricted version of the eGII (Table B-5 in the Appendix).

[16] Moscona *et al.* [2018] show how Murdock boundaries reflect borders of ethnic settlements today.

7.2. Sexual Violence and Cultural Distance in Gender Norms

To explore whether the prevalence and intensity of war-related sexual violence are explained by the cultural distance in gender norms of the combatants, we take advantage of the dyadic structure of the data.[17] For each actor in every year of conflict, we have information on their ethnic characteristics, on their use of sexual violence and, most importantly, on their opponents. Two actors fighting against each other in a conflict constitute a *dyad*. For instance, the government of Chad (corresponding to the "Zaghawa, Bideyat" ethnic group) and the *Comité de Sursaut National pour la Paix et la Démocratie* (CSNPD) rebel group (corresponding to the "Sara" ethnic group) form a dyad in our dataset. Since we have information on sexual violence perpetrated by both actors, our dyadic dataset is bidirectional. Therefore, every dyad involved in a conflict event appears twice in the dataset. In one instance, the government of Chad is the perpetrator and the CSNPD is the victim. The variable $SVAC$, in this case, captures the intensity of sexual violence inflicted by the government of Chad to the CSNPD. In a second instance, the government of Chad is the victim, and the CSNPD is the perpetrator. In this case, the variable $SVAC$ captures the intensity of sexual violence inflicted by the rebel group CSNPD to the government of Chad.

We first construct a measure of absolute distance between the perpetrator's gender inequality index and the victim's gender inequality index as follows:

$$\text{eGII}_{pv}^{Dist} = |\text{eGII}_p - \text{eGII}_v|$$

Next, we estimate the following specification, in the spirit of a gravity approach (see Grosjean [2011] and Serafinelli and Tabellini [2017] as examples of gravity equations applied to culture):

$$\text{SVAC}_{pvct} = \alpha + \gamma \text{eGII}_{pv}^{Dist} + \Phi_c + \tau_t + \omega_c t + P_p + \epsilon_{pvct} \tag{3}$$

The dependent variable is an index capturing the intensity of sexual violence perpetrated by actor p against actor v during conflict c in year t; Φ_c and τ_t denote conflict and year fixed effect, respectively, and $\omega_c t$ is a conflict-specific time trend. The inclusion of perpetrator fixed effect (P_p) allows to control for perpetrator-specific time invariant characteristics such as own gender inequality index, overall aggressiveness, and other ethnic traits. Standard errors are clustered at the dyadic level. We restrict the sample to inter-ethnic conflicts only, i.e., to those where we are able to assign distinct ethnic identities to the perpetrator and the victim.

[17] Table 1 reports an extract of our dataset.

To further explore the nature of the cultural clash, we split the absolute distance measure into two components, and separately assess the impact of (i) the perpetrator being *more gender unequal* than the victim and (ii) the perpetrator being *less gender unequal* than the victim when explaining the use of sexual violence in a conflict:

$$\text{Perpetrator More Unequal}_{pv} = \begin{cases} |\text{eGII}_p - \text{eGII}_v| & \text{if eGII}_p > \text{eGII}_v \\ 0 & \text{otherwise.} \end{cases}$$

$$\text{Perpetrator Less Unequal}_{pv} = \begin{cases} |\text{eGII}_p - \text{eGII}_v| & \text{if eGII}_p < \text{eGII}_v \\ 0 & \text{otherwise.} \end{cases}$$

To tease out the separate effect of these two distinct components of cultural distance, we estimate the following:

$$\begin{aligned} \text{SVAC}_{pvct} = \alpha + \eta_1 &\text{Perpetrator More Unequal}_{pv} \\ &+ \eta_2 \text{Perpetrator Less Unequal}_{pv} + \Phi_c + \tau_t + \omega_c t + P_p + \epsilon_{pvct} \end{aligned} \quad (4)$$

This specification is equivalent to the one in Equation 3, and it differs only in that it substitutes the cultural distance measure with its two main components. In this specification, η_1 and η_2 separately capture the effects of two distinct cultural clashes: one where the perpetrator faces an opponent characterized by more gender-equal cultural norms compared to its own norms (η_1), and one where the perpetrator is confronted with an opponent characterized by more gender-unequal cultural norms (η_2).

7.3. Results

Results are reported in Table 8. Column 1 presents the coefficient estimate of γ in Equation 3. There is a positive and significant association between the absolute cultural distance in gender norms of the combatants and the intensity in sexual violence in conflict. One standard deviation increase in the absolute distance in the eGII of the combatants increases the intensity of sexual violence by 0.21 standard deviations.

Columns (2)-(5) unpack this association, and separately assess the role played by the perpetrator's own eGII and by two distinct cultural clashes: when (i) the perpetrator is confronted with an opponent who holds more gender-equal cultural norms (*Perpetrator more unequal*) (ii) the perpetrator is confronted with an opponent that holds more gender-unequal norms (*Perpetrator less unequal*). Column (2) displays results of a horse-race between the perpetrator's eGII and the absolute cultural distance when the perpetrator is more gender unequal than the victim. The coefficient on eGII is positive, but statistically

insignificant and smaller than in Table 6, which showed the association between the eGII alone and sexual violence. Instead, the coefficient on *Perpetrator more unequal* is larger in magnitude, and significant at the 10 percent level. Column 3 includes *Perpetrator less unequal*, i.e., the absolute cultural distance when the perpetrator holds more equitable gender norms than the victim. The latter factor seems to not be positively associated with the use of sexual violence: the coefficient is small and not significantly different from zero. Instead, the coefficient on the perpetrator's eGII is large and significant in this specification.

Column (4) shows coefficient estimates for a specification that simultaneously includes the perpetrator's eGII and the two different cultural clashes. All coefficients estimates are positive, but the largest and only significant one is that on the perpetrator being more gender-unequal than the victim. This coefficient can be interpreted as follows: when a gender-unequal perpetrator with a eGII of 1 faces a gender-egalitarian victim with a eGII of 0, sexual violence intensity is 1.51 higher than when the perpetrator and the victim hold the same gender norms. In column (5), instead of controlling for the perpetrator's eGII, we add perpetrator fixed effects. Crucially, this allows us to account for any time-invariant perpetrator's characteristics that may confound the results, like other ethnic traits (including the perpetrator's own eGII), overall aggressiveness, whether the perpetrator is a state force or a rebel group. The coefficient barely changes in magnitude, and its statistical significance increases.

7.4. Interpretation of the Results

How can we explain the positive and significant coefficient of *Perpetrator more unequal* in Table 8? One potential explanation is that perpetrators may strategically use sexual violence to target a valuable asset in the opponent's society (i.e., women). This behavior would arise when women have a prominent role on the victim's society—and consequently, when perpetrators are more likely to be more gender-unequal. This could occur in the absence of a cultural clash driven by divergent gender norms. To better understand the underlying mechanisms that trigger sexual violence, we re-run our main specification controlling for victim's characteristics, that take into account the economic value of women in victim's society—as well as other time-invariant ethnic characteristics. By and large, our results are robust to this procedure (see Table B-7). When controlling for the victim's eGII rather than the perpetrator's, the coefficient on *Perpetrator more unequal* stays large and significant. When adding victim's fixed effects, the same coefficient loses significance despite maintaining a similar magnitude.

Another potential explanation for our results is that women in some ethnic groups might be more vulnerable and therefore more likely to be victimized, independently of the cultural distance between the two combatants. Again, the results in Table B-7, controlling for victim's characteristics, can help us to shed light on this alternative channel. When comparing Table 8 with Table B-7 we do find that the magnitude of *Perpetrator more unequal* is larger and the magnitude of *Perpetrator less unequal* is smaller and negative in most cases. This is consistent with the former coefficient being potentially biased downwards and the latter biased upwards when we do not account for the fact that women in the rival's side might be less and more unprotected and easily targeted, respectively.[18] However, we still find that cultural distance matters for explaining sexual violence during conflict, even after controlling for perpetrator and victim-specific traits.

A subsequent natural question to ask would be, what is the mechanism by which cultural distance triggers sexual violence? Why do we only find results when looking at a certain cultural clash, that is, when the perpetrator is more gender-unequal than the victim? The asymmetry of our findings is consistent with an explanation based on the identity model by Akerlof and Kranton (2000). Perpetrators that confront more gender-equal opponents might perceive that the relatively better position of women is a threat to their own norms or ideals, and might resort to sexual violence to reinstate the disutility suffered by this encounter. However, perpetrators that face a more gender-unequal opponent might not feel that their masculinity is threatened by the worse relative position of women in the opponents' societies. Since their identity utility might not be affected, they are less likely to respond with sexual violence to the distance in cultural norms.

To further validate our results, in the next Subsections we conduct a battery of robustness checks, and explore the relationship between gender-norms distance and general (i.e., non-sexual) violence as well as the relationship between other measures of cultural distance—linguistic and religious—and sexual violence in conflict.

7.5. Robustness Checks

We report robustness checks in Tables B-8 to B-11 in the Appendix. Given that our measure of cultural distance in gender norms is time-invariant, we show that our results are similar when running the same specifications abstracting from the temporal variation in the data (Table B-8).

Table B-9 shows that the coefficients are robust to the inclusion of alternative sets of

[18] We are assuming that women are less vulnerable in a more gender-equal society.

fixed effects, and to alternative versions of the eGII. In column (1), we include conflict-year fixed effects, to account for any conflict-year specific factor that may confound the results (e.g., how cruel the conflict was in that specific year, changes in military tactics during the conflict, and so on). In column (2), we add country fixed effects. Our results remain unchanged when using the unweighted version of the eGII, and when using the restricted version of the eGII that includes only five ethnic characteristics.

We also test whether our results hold when assigning an alternative measure of the eGII to state military forces. This exercise is motivated by the fact that the composition of the state military may not reflect the ethnic identity of the government. For instance, it is possible that the election of a new government does not (at least immediately) result into an alignment of the military with the ethnic identity of the new ethnic groups in power. For this reason, we explore whether our results hold when assigning to the government a more conservative measure, i.e. the average eGII of all the Murdock ethnic groups within a country, weighted by the size of each group's settlement area. Table B-10 shows that replacing the government's eGII with the country average leaves the results almost unchanged: the *Perpetrator more unequal* coefficient stays significant and slightly increases in magnitude compared to our baseline specifications in Table 8. Conversely, the *Perpetrator less unequal* coefficient remains insignificant, and its size slightly decreases.

Finally, in Table B-11, we show that results are robust to multi-way clustering, i.e. to clustering standard errors at the level of the first and of the second actor in a given pair. This allows for arbitrary correlations of the error term within a group of actors pairs that share the same perpetrator or that share the same victim.

7.6. Gender-Norms Clash and General Violence

Cultural distance in gender norms between the combatants may be correlated to other dissimilarities between ethnic groups, which could in turn generate more violent conflicts, or more violent episodes within a conflict. If this was the case, sexual violence would only be a byproduct of general violence, and the cultural clash we are estimating would not be specifically linked to a gender-based form of violence.

The inclusion of conflict fixed effects in our main specification partially alleviates this concern because it accounts for the overall cruelty in a conflict. However, to fully rule out this alternative explanation, we run a placebo test using a different measure of violence as the outcome variable. We exploit information on the number of fatalities experienced by a conflict actor in every year of conflict to construct an index similar to the sexual violence

variable, ranging between 0 to 3 depending on the number of recorded deaths. Since this measure of general violence is also bidirectional, we can run the same specifications of Equations 2 and 3 having as an outcome variable the number of deaths inflicted by the perpetrator on the victims.

Table 9 reports the results of this exercise. Column (1) shows that the distance in gender norms of the combatants is not associated with the number of deaths inflicted to the victims: the coefficient is small and not statistically different from zero. The same holds for the coefficients in columns (2) to (5). Neither the perpetrator's own gender norms, nor the cultural clashes are positively and significantly associated with general violence. If at all, these elements seem to be negatively associated with general violence, although none of these coefficients is statistically significant. Taken together, this suggests that the perpetrator's eGII and the cultural clash in gender norms between perpetrator and victim explain only the use of *gender-based* forms of violence.

7.7. Distance in Gender Norms and Other Measures of Cultural Distance

In this section, we assess whether combatants' clashes in gender norms are the main driver of the use of sexual violence, or whether clashes in other cultural dimensions are similarly powerful in explaining this phenomenon. Based on existing work, Mokyr [2016] has recently defined culture as a "set of beliefs, values, and preferences, capable of affecting behavior, that are socially [...] transmitted and that are shared by a subset of society". Gender norms are part of the beliefs, values, and preferences that constitute culture. Therefore, while we found that distance in gender norms is associated with the use and intensity of sexual violence, we are interested in exploring whether differences in other cultural traits unrelated to gender are an equally important factor. To disentangle the role of gender norms from other aspects of culture, we exploit linguistic and religious distance, widely-used proxies for cultural differences between populations.[19]

The use of linguistic distance is motivated by the notion that language is a salient dimension of culture, which is transmitted through generations (Spolaore and Wacziarg [2016]). Different languages are the result of horizontal separations between populations, and these separations are likely to go hand in hand with cultural divergence. We use Fearon's [2003] measure of linguistic distance (called cladistic distance), which is based

[19]Spolaore and Wacziarg [2016] use other country-level measures of cultural heterogeneity, like answers to World Value Survey (WVS) questions. We abstain from using the WVS due to potential reverse causality between this measure and our outcome variable. In addition, Spolaore and Wacziarg [2016] proposed genetic distance as a summary measure for populations' *relatedness*. This measure is unfeasible in our context due to a too wide categorization of ethnic groups in the original source of genetic data.

on linguistic trees in the Ethnologue, a comprehensive database of more than 7,000 known living languages. We merge information on languages spoken by ethnic groups through the Ethnic Power Relations-Ethnic Dimensions (EPR-ED) dataset, and compute distances between each pair of languages based on the number of common nodes in the tree. This allows us to compute a measure of linguistic distance between ethnic groups, and ultimately, between the perpetrator and the victim. Following the same methodology, we construct a measure of religious distance. Sections A-2 and A-3 in the Appendix provide additional details on how we construct these measures.

Figure 15 plots the correlation between linguistic distance between the combatants and our measure of distance in gender norms. Not surprisingly, the correlation is positive, suggesting that conflict actors that are linguistically distant are on average also more likely to differ in their gender norms. However, the figure and the respective correlation coefficient (0.25, statistically significant at the one percent level) also suggest that distance in gender norms is not a perfect predictor of overall cultural distance. On the other hand, distance in gender norms is uncorrelated with religious distance (see Figure 16).[20] Taken together, this suggests that our gender norms measure captures a dimension distinct from already-proposed aspects of cultural distance.

To assess whether cultural clashes in traits unrelated to gender norms are associated with sexual violence, we first regress linguistic distance on distance in gender norms, and obtain residuals of this regression. These residuals capture the component of cultural distance that is unexplained by differences in gender norms. We re-run our main specifications controlling for this component of cultural distance and, alternatively, for linguistic distance. As can be seen in columns (2)-(3) and columns (5)-(6), this leaves the coefficients in our main specifications almost unchanged. Columns (7) and (8) in Table 10 show that neither overall cultural distance nor residuals alone can explain actors' use of sexual violence in conflict (the coefficient is small in magnitude and insignificant).

Table 11 repeats the same exercise using religious distance as a proxy for cultural distance between perpetrator and victim. Again, controlling for residuals or religious distance does not affect our main results. Interestingly, column (8) shows that, if at all, religious distance is *negatively* associated with the use of sexual violence.[21] Taken together, these results suggest that what matters in explaining the use of sexual violence is not cultural distance in general, but a specific clash in cultural norms related to gender.

[20] Religious distance and linguistic distance are positively correlated. See Figure B-5 in the Appendix.

[21] Given that distance in gender norms and religious distance are uncorrelated, it is not surprising that the coefficient in column 7 is almost identical to the one in column 8.

8. Conclusion

Why do some conflict actors systematically rape, while others never do so? In this paper, we advance and test a new hypothesis for the use of sexual violence in armed conflict.

We find that armed actors characterized by more gender-unequal norms are more likely to engage in sexual violence during ethnic conflict. However, we show that this explanation for sexual violence—stemming from gender inequality on the perpetrator's side—is incomplete. The prevalence and the intensity of war-related sexual violence is better explained when considering both the perpetrator's and the victim's gender norms. In particular, sexual violence emerges and intensifies when there is a *clash of conceptions* between combatants on what is the appropriate role of men and women in society. Cultural distance in gender norms between perpetrator and victim explains sexual violence more strongly than the perpetrator's own gender inequality.

When examining this relationship further, we uncover that the effect is driven by a specific cultural clash, i.e., when the perpetrator holds more gender-unequal norms than the victim. We show that this just-described pattern is specific to gender-based violence, and that it does not explain the intensity of general violence perpetrated by an armed actor, measured by deaths inflicted on the opponent. Moreover, conflict-related sexual violence is not driven by general cultural differences, but by differences in gender norms.

Our contribution in this paper is threefold. First, we enrich existing conflict data sources with ethnic characteristics of the groups involved. This novel dataset may constitute a potentially valuable resource for future contributions in the conflict literature. While the literature mas mainly focused on understanding how ethnic diversity or cultural distance determine the onset of conflicts, we show that they can also explain how violence manifests once the conflict has started.

Second, we propose and validate an ethnic gender inequality index at the ethnic-group level for Africa. This index complements the gender inequality index (GII) introduced in 2010 by the United Nations Development Programme, which is constructed at the country level and based on contemporary variables (reproductive health, empowerment and labor market participation). Our Gender Inequality Index is instead constructed at the ethnic group level, is based on anthropological notions of gender (in)equality, and aims at capturing the deeply entrenched norms of a society.

Finally, in line with recent literature on the cultural determinants of intimate-partner violence (Alesina *et al.* [forthcoming]; Tur-Prats [2019]), we find that violence against women during wartime shares the same fundamental causes as violence against women

during peacetime. From a policy perspective, this suggests that policies aimed at changing gender norms might have an effect on all the different manifestations of violence against women.

References

Alesina, A., Brioschi, B. and La Ferrara, E. (forthcoming). Violence Against Women: A Cross-cultural Analysis for Africa. *Economica*.

—, Giuliano, P. and Nunn, N. (2013). On the Origins of Gender Roles: Women and the Plough. *The Quarterly Journal of Economics*, **128** (2), 469–530.

Arbatli, C. E., Ashraf, Q. H. and Galor, O. (2013). The nature of civil conflict. *Available at SSRN 2364127*.

Ba, I. and Bhopal, R. (2017). Physical, mental and social consequences in civilians who have experienced war-related sexual violence: a systematic review (1981–2014). *Public Health*, **142**, 121–135.

Bahgat, K., Nordås, R. and Østby, G. (2016). Geocoded SVAC Dataset (GEO-SVAC) version 1.0. Peace Research Institute Oslo (PRIO).

Becker, A. (2018). On the Economic Origins of Constraints on Women's Sexuality.

—, Enke, B. and Falk, A. (2018). *Ancient origins of the global variation in economic preferences*. Tech. rep., National Bureau of Economic Research.

Becker, G. S. (1957). 1971. the economics of discrimination.

Blattman, C. and Miguel, E. (2010). Civil war. *Journal of Economic literature*, **48** (1), 3–57.

Boserup, E. (1970). *Woman's role in economic development*.

Bremer, S. A. (2000). Who fights whom, when, where, and why? *What do we know about war*, pp. 23–36.

Caselli, F. and Coleman, W. J. (2013). On the theory of ethnic conflict. *Journal of the European Economic Association*, **11** (suppl_1), 161–192.

Cederman, L.-E., Min, B. and Wimmer, A. (2012). Ethnic Armed Conflict dataset.

Cohen, D. K. (2013). Explaining Rape During Civil War: Cross-National Evidence (1980–2009). *American Political Science Review*, **107** (3), 461–477.

— and Nordås, R. (2014). Sexual Violence in Armed Conflict: Introducing the SVAC Dataset, 1989–2009. *Journal of Peace Research*, **51** (3), 418–428.

Croicu, M. and Sundberg, R. (2017). UCDP GED Codebook version 18.1. Department of Peace and Conflict Research, Uppsala University.

Desmet, K., Ortuño-Ortín, I. and Wacziarg, R. (2017). Culture, ethnicity, and diversity. *American Economic Review*, **107** (9), 2479–2513.

Fearon, J. D. (2003). Ethnic and cultural diversity by country. *Journal of economic growth*, **8** (2), 195–222.

Friedl, E. (1978). Society and sex roles. *Human Nature*, **1** (4), 68–75.

Girardin, L., Hunziker, P., Cederman, L.-E., Bormann, N.-C. and Vogt, M. (2015). Growup-geographical research on war, unified platform. *ETH Zurich. http://growup. ethz. ch*.

Gleditsch, N. P., Wallensteen, P., Eriksson, M., Sollenberg, M. and Strand, H. (2002). Armed Conflict 1946-2001: A New Dataset. *Journal of peace research*, **39** (5), 615–637.

Gneezy, U., Leonard, K. L. and List, J. A. (2009). Gender differences in competition: Evidence from a matrilineal and a patriarchal society. *Econometrica*, **77** (5), 1637–1664.

Grosjean, P. (2011). The weight of history on european cultural integration: a gravity approach. *American Economic Review*, **101** (3), 504–08.

Guiso, L., Sapienza, P. and Zingales, L. (2016). Long-term persistence. *Journal of the European Economic Association*, **14** (6), 1401–1436.

Holsoe, S. E. and Lauer, J. J. (1976). Who are the kran/guere and the gio/yacouba? ethnic identifications along the liberia-ivory coast border. *African Studies Review*, **19** (1), 139–150.

Huntington, S. P. (2000). The clash of civilizations? In *Culture and Politics*, Springer, pp. 99–118.

International Criminal Court (2000). Elements of crimes. PCNICC /2000/1/Add.2.

International Criminal Court (2002). Rome Statute. Document A/CONF.183/9.

Korotayev, A. (2003). Division of labor by gender and postmarital residence in cross-cultural perspective: A reconsideration. *Cross-Cultural Research*, **37** (4), 335–372.

Lowes, S. (2017). Matrilineal kinship and spousal cooperation: Evidence from the matrilineal belt. *Unpublished manuscript. URL: https://scholar. harvard. edu/files/slowes/files/lowes_matrilineal. pdf*.

Martin, M. K. and Voorhies, B. (1975). *Female of the Species*. Columbia University Press.

Meger, S. (2016). *Rape loot pillage: The political economy of sexual violence in armed conflict*. Oxford University Press.

Michalopoulos, S. and Papaioannou, E. (2016). The long-run effects of the scramble for africa. *American Economic Review*, **106** (7), 1802–48.

Mokyr, J. (2016). *A culture of growth: the origins of the modern economy*. Princeton University Press.

Moscona, J., Nunn, N. and Robinson, J. A. (2018). *Social structure and conflict: Evidence from sub-saharan africa*. Tech. rep., National Bureau of Economic Research.

Murdock, G. P. (1967). Ethnographic Atlas.

Nunn, N. and Wantchekon, L. (2011). The slave Trade and the Origins of Mistrust in Africa. *The American Economic Review*, **101** (7), 3221–3252.

Putterman, L. and Weil, D. N. (2010). Post-1500 population flows and the long-run determinants of economic growth and inequality. *The Quarterly journal of economics*, **125** (4), 1627–1682.

Sanday, P. R. (1981a). *Female power and male dominance: On the origins of sexual inequality*. Cambridge University Press.

— (1981b). The socio-cultural context of rape: A cross-cultural study. *Journal of social issues*, **37** (4), 5–27.

Schneider, D. M. and Gough, K. (1961). *Matrilineal kinship*. Univ of California Press.

Serafinelli, M. and Tabellini, G. (2017). Creativity over time and space. *Available at SSRN 3070203*.

Skjelsbaek, I. (2001). Sexual Violence and War: Mapping Out a Complex Relationship. *European Journal of International Relations*, **7** (2), 211–237.

Spolaore, E. and Wacziarg, R. (2009). The diffusion of development. *The Quarterly journal of economics*, **124** (2), 469–529.

— and — (2016). Ancestry, language and culture. In *The Palgrave handbook of economics and language*, Springer, pp. 174–211.

Sundberg, R. and Melander, E. (2013). Introducing the UCDP georeferenced event dataset. *Journal of Peace Research*, **50** (4), 523–532.

Symons, D. (1979). The evolution of human sexuality.

Taylor, C. C. (1999). A gendered genocide: Tutsi women and hutu extremists in the 1994 rwanda genocide. *PoLAR*, **22**, 42.

Teso, E. (2018). The long-term effect of demographic shocks on the evolution of gender roles: Evidence from the transatlantic slave trade. *Journal of the European Economic Association*, **17** (2), 497–534.

Thornhill, R. and Palmer, C. T. (2001). *A Natural History of Rape: Biological Bases of Sexual Coercion*. MIT press.

Tur-Prats, A. (2019). Family types and intimate partner violence: A historical perspective. *Review of Economics and Statistics*, **101** (5), 878–891.

Vogt, M., Bormann, N.-C., Rüegger, S., Cederman, L.-E., Hunziker, P. and Girardin, L. (2015). Integrating Data on Ethnicity, Geography, and Conflict: The Ethnic Power Relations Data Set Family. *Journal of Conflict Resolution*, **59** (7), 1327–1342.

Voigtländer, N. and Voth, H.-J. (2012). Persecution perpetuated: the medieval origins of anti-semitic violence in nazi germany. *The Quarterly Journal of Economics*, **127** (3), 1339–1392.

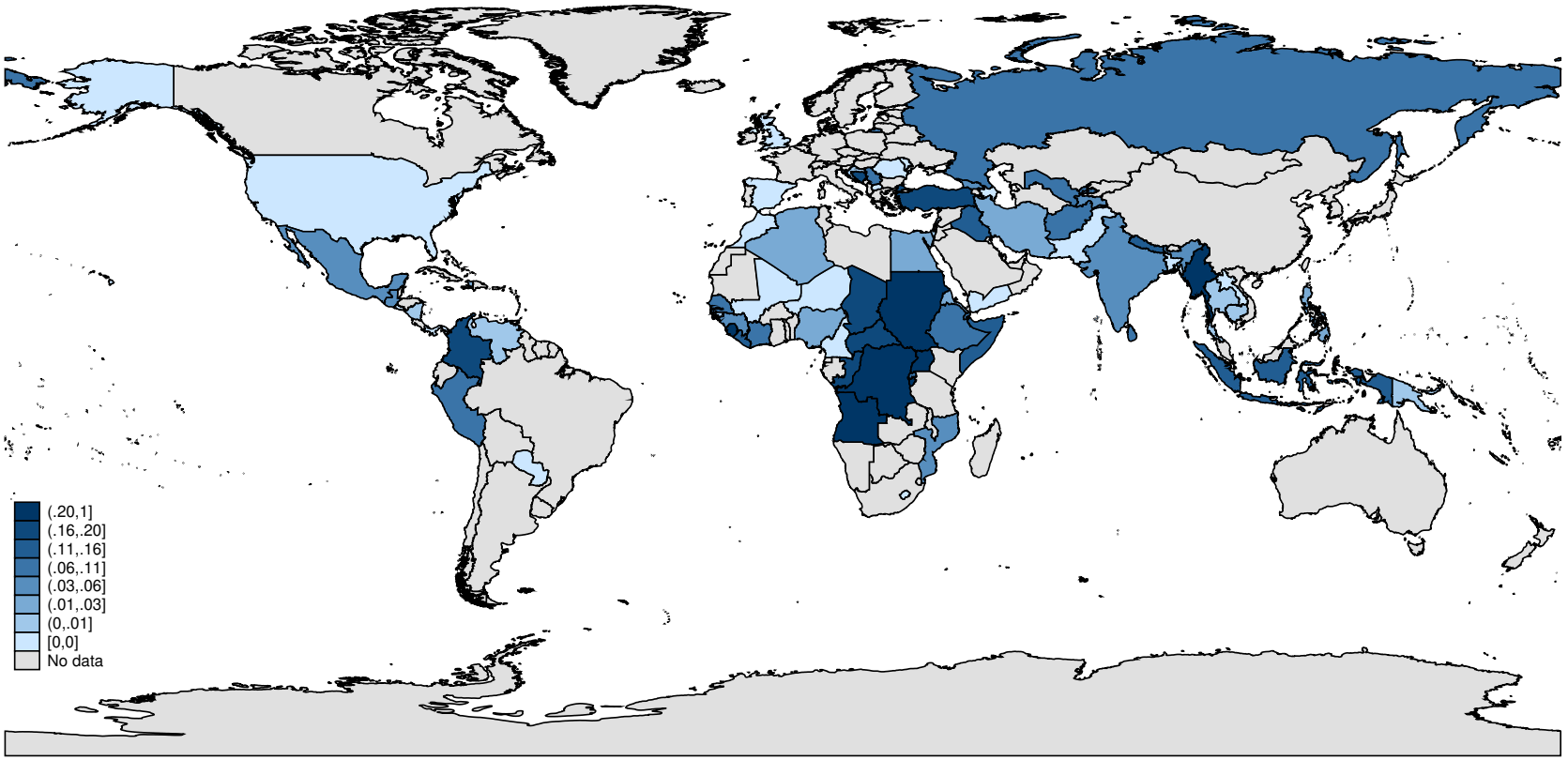
Wimmer, A., Cederman, L.-E. and Min, B. (2009). Ethnic politics and armed conflict: A configurational analysis of a new global data set. *American Sociological Review*, **74** (2), 316–337.

Wood, E. J. (2009). Armed Groups and Sexual Violence: When is Wartime Rape Rare? *Politics & Society*, **37** (1), 131–161.

Wucherpfennig, J., Metternich, N. W., Cederman, L.-E. and Gleditsch, K. S. (2012). Ethnicity, the state, and the duration of civil war. *World Politics*, **64** (1), 79–115.

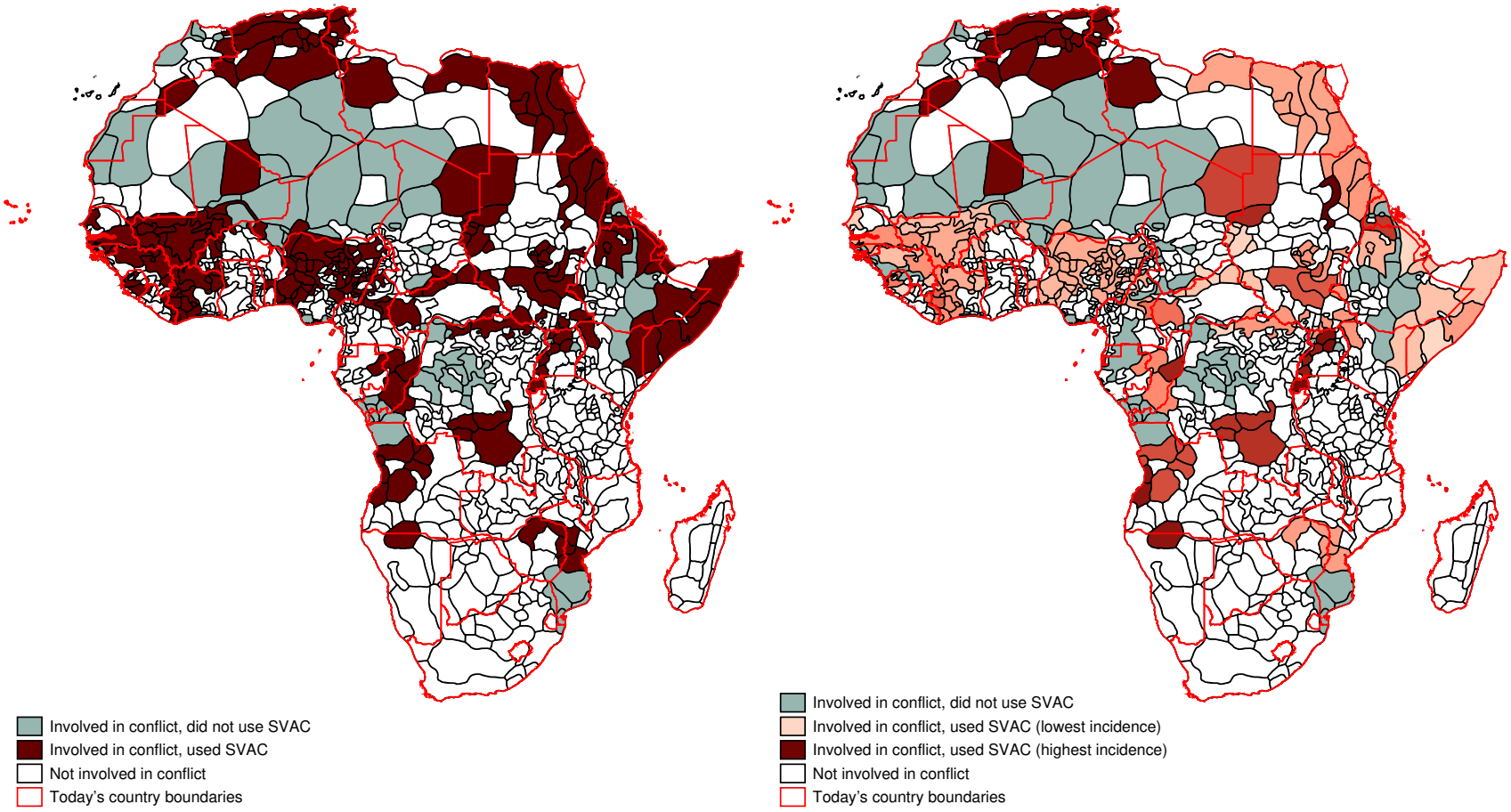
Figures and Tables

FIGURE 1: *Global prevalence of sexual violence in armed conflict (1989-2009) according to US State Department*



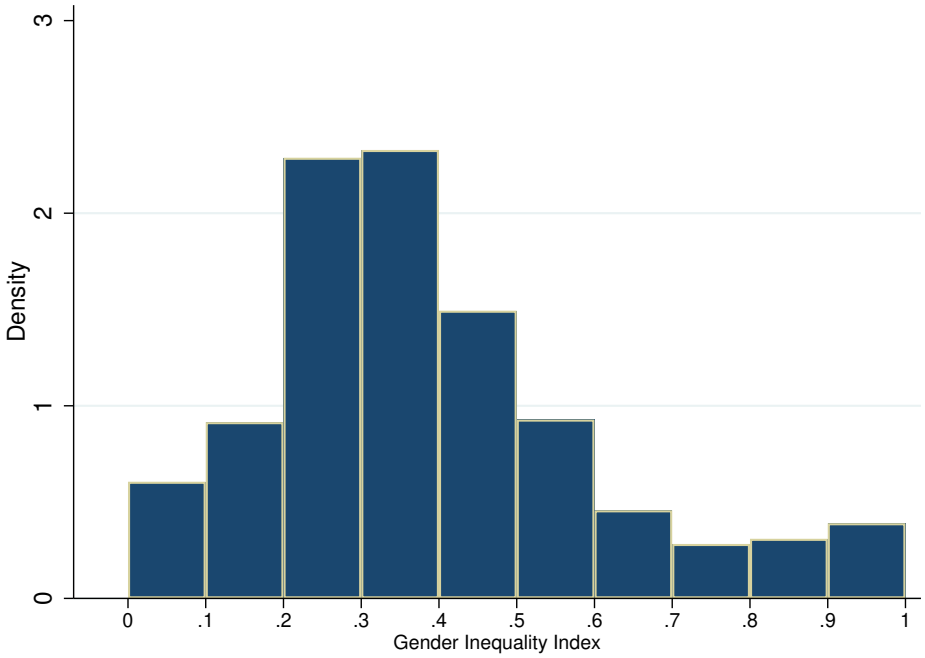
NOTES: The map reports an index ranging between 0 and 1 capturing the total incidence of sexual violence in armed conflict for the period 1989-2009 worldwide as reported by the US state department. Source: SVAC dataset.

FIGURE 2: *Murdock ethnic groups' use of sexual violence in armed conflict (1989-2009) in Africa*



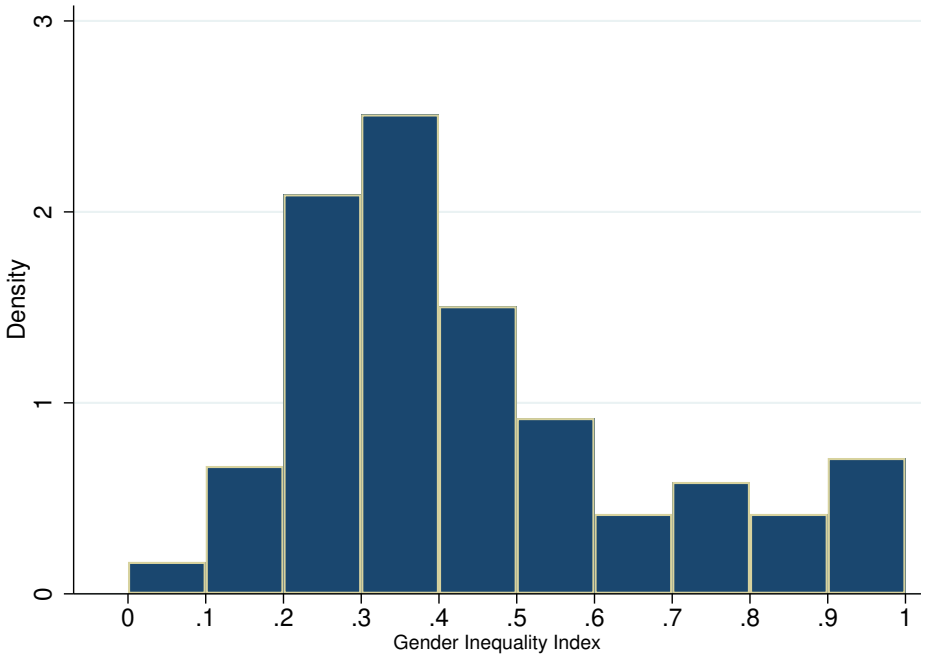
NOTES: Left: involvement in conflict and use of sexual violence by Murdock ethnic groups in Africa. Right: involvement in conflict and incidence of sexual violence used by Murdock's ethnic groups in Africa, through an index varying between 0-1, which captures the total incidence of sexual violence in armed conflict for the period 1989-2009. Sources: Murdock Ethnographic Atlas and GEO-SVAC dataset.

FIGURE 3: *Distribution of the eGII in Africa*



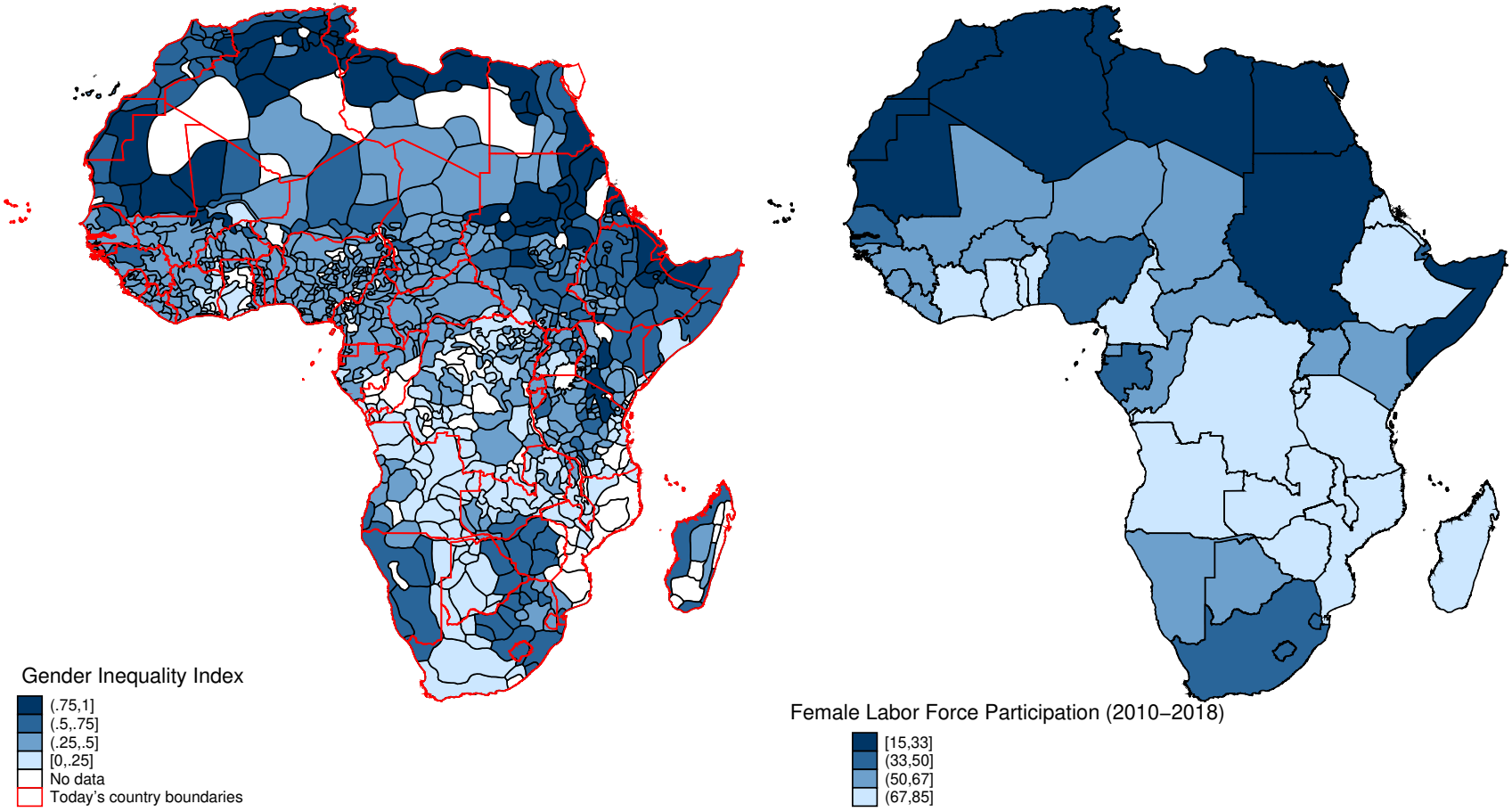
NOTES: Distribution of the eGII in Africa. Mean (standard deviation) of the index: 0.40 (.21).

FIGURE 4: *Distribution of the eGII in our sample*



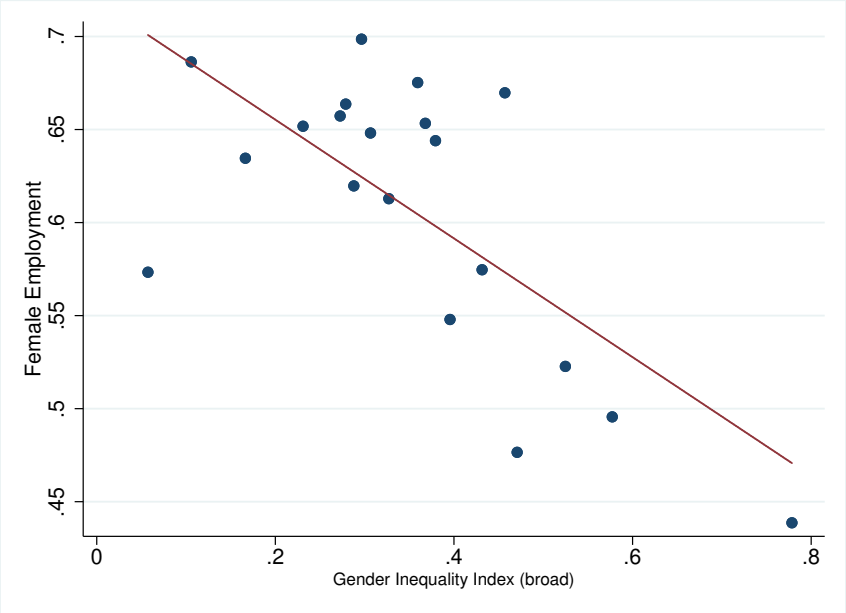
NOTES: Distribution of the eGII in our sample. Mean (standard deviation) of the index: 0.45 (.22).

FIGURE 5: *Distribution of the eGII across Africa and female labor force participation*



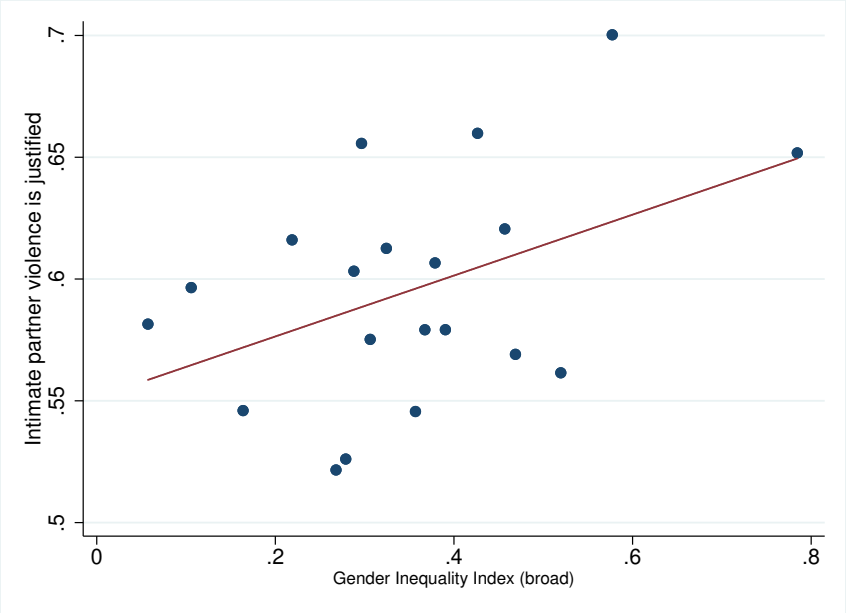
NOTES: Distribution of the eGII across Murdock's ethnicities in Africa and contemporary country borders. Right: Female labor force participation at the country level (2010-2018) for women older than 15. Darker colors denote lower participation. Source: International Labor Organization.

FIGURE 6: *Correlation between the eGII and female employment in sub-Saharan Africa (ethnicity level)*



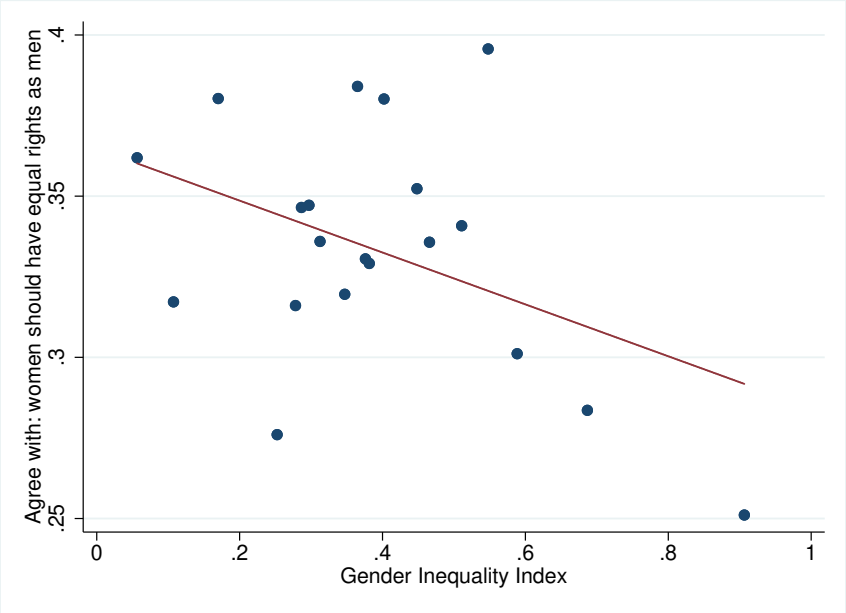
NOTES: Correlation between the eGII and female employment in sub-Saharan Africa (32 countries) at the Murdock ethnicity level. Correlation coefficient: -0.28*** Source: Ethnographic Atlas and Demographic and Health (DHS) survey.

FIGURE 7: *Correlation between the eGII and justification of intimate partner violence in sub-Saharan Africa*



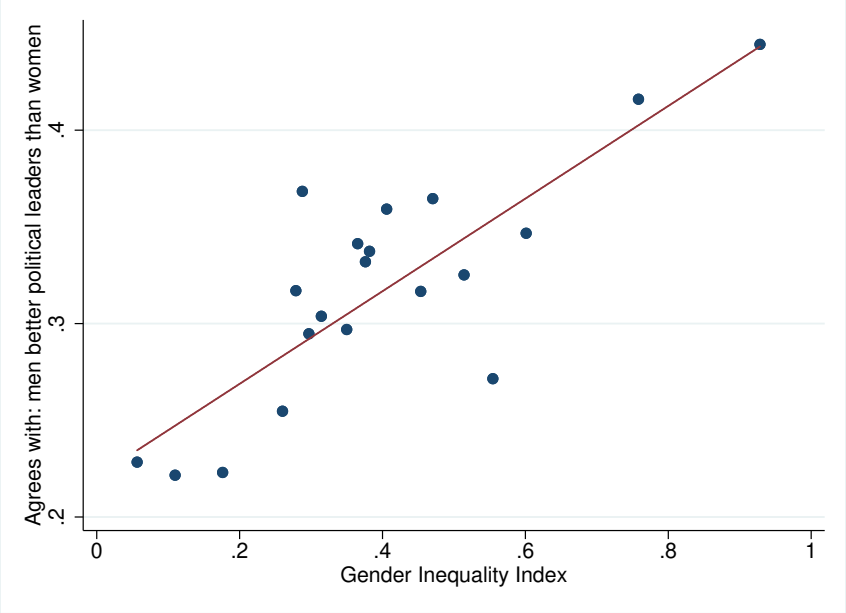
NOTES: Correlation between the eGII and justification of intimate partner violence in sub-Saharan Africa (32 countries) at the Murdock ethnicity level. Correlation coefficient: -0.11** Source: Ethnographic Atlas and Demographic and Health (DHS) survey.

FIGURE 8: *Correlation between the eGII and agreeing with the statement: "Women should have equal rights and receive the same treatment as men do"*



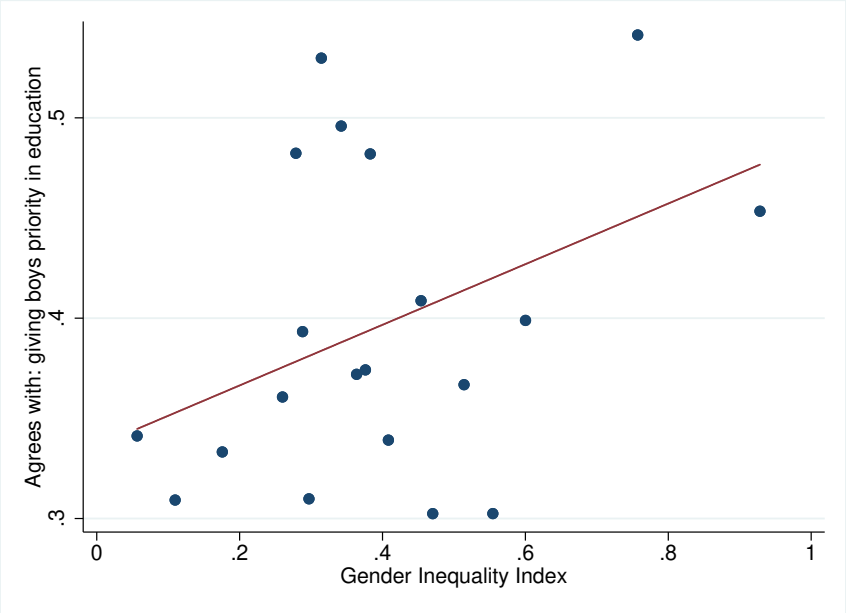
NOTES: Correlation between the eGII and agreement with the statement: "Women should have equal rights and receive the same treatment as men do" in Africa (35 countries) at the Murdock ethnicity level. Correlation coefficient: -0.10** Sources: Ethnographic Atlas and Afrobarometer.

FIGURE 9: *Correlation between the eGII and agreeing with the statement: "Men make better political leaders than women"*



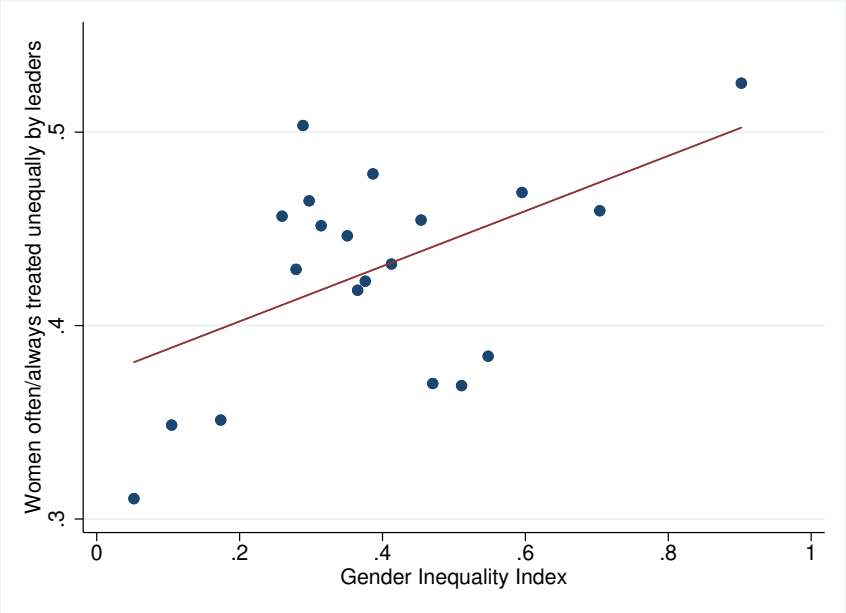
NOTES: Correlation between the eGII and agreement with the statement: "Men make better political leadersthan women, and should be elected rather than women" in Africa (33 countries) at the Murdock ethnicity level. Correlation coefficient: 0.31*** Sources: Ethnographic Atlas and Afrobarometer.

FIGURE 10: *Correlation between the eGII and agreeing with the statement: "If funds for schooling are limited, a boy should always receive an education in school before a girl."*



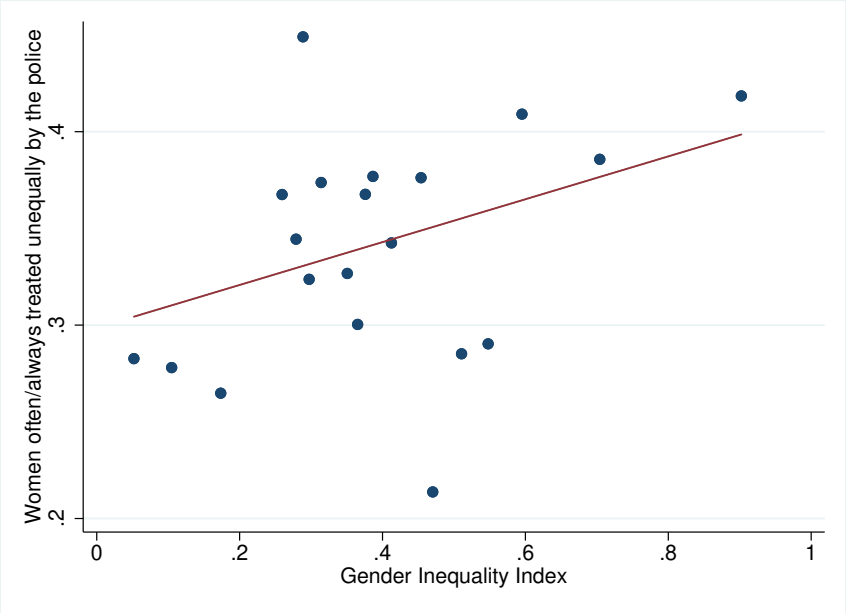
NOTES: Correlation between the eGII and agreement with the statement: "If funds for schooling are limited, a boy should always receive an education in school before a girl" in Africa (33 countries) at the Murdock ethnicity level. Correlation coefficient: 0.12** Sources: Ethnographic Atlas and Afrobarometer.

FIGURE 11: *Correlation between the eGII and reporting that "women are often/always treated unequally by traditional leaders"*



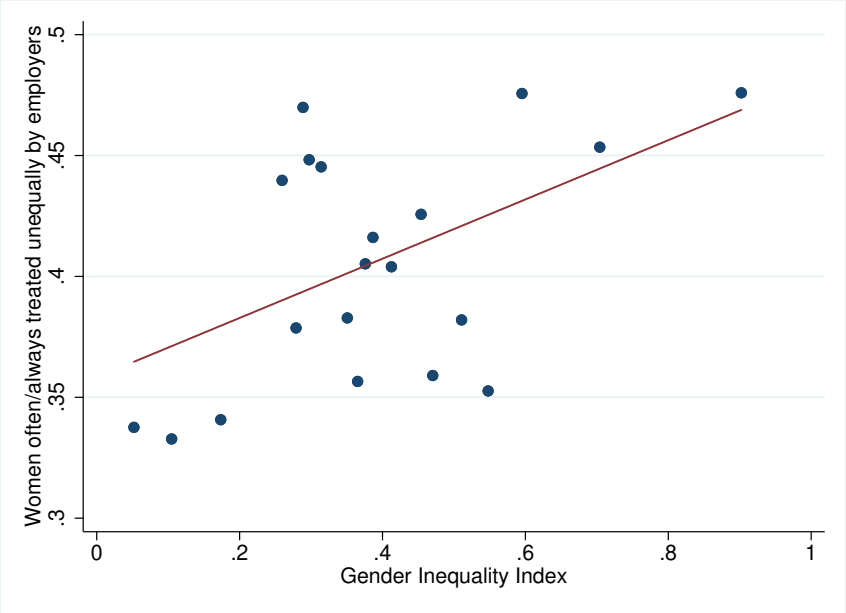
NOTES: Correlation between the eGII and reporting that "women are often/always treated unequally by traditional leaders" in Africa (32 countries) at the Murdock ethnicity level. Correlation coefficient: 0.15*** Sources: Ethnographic Atlas and Afrobarometer.

FIGURE 12: *Correlation between the eGII and reporting that "women are often/always treated unequally by the police"*



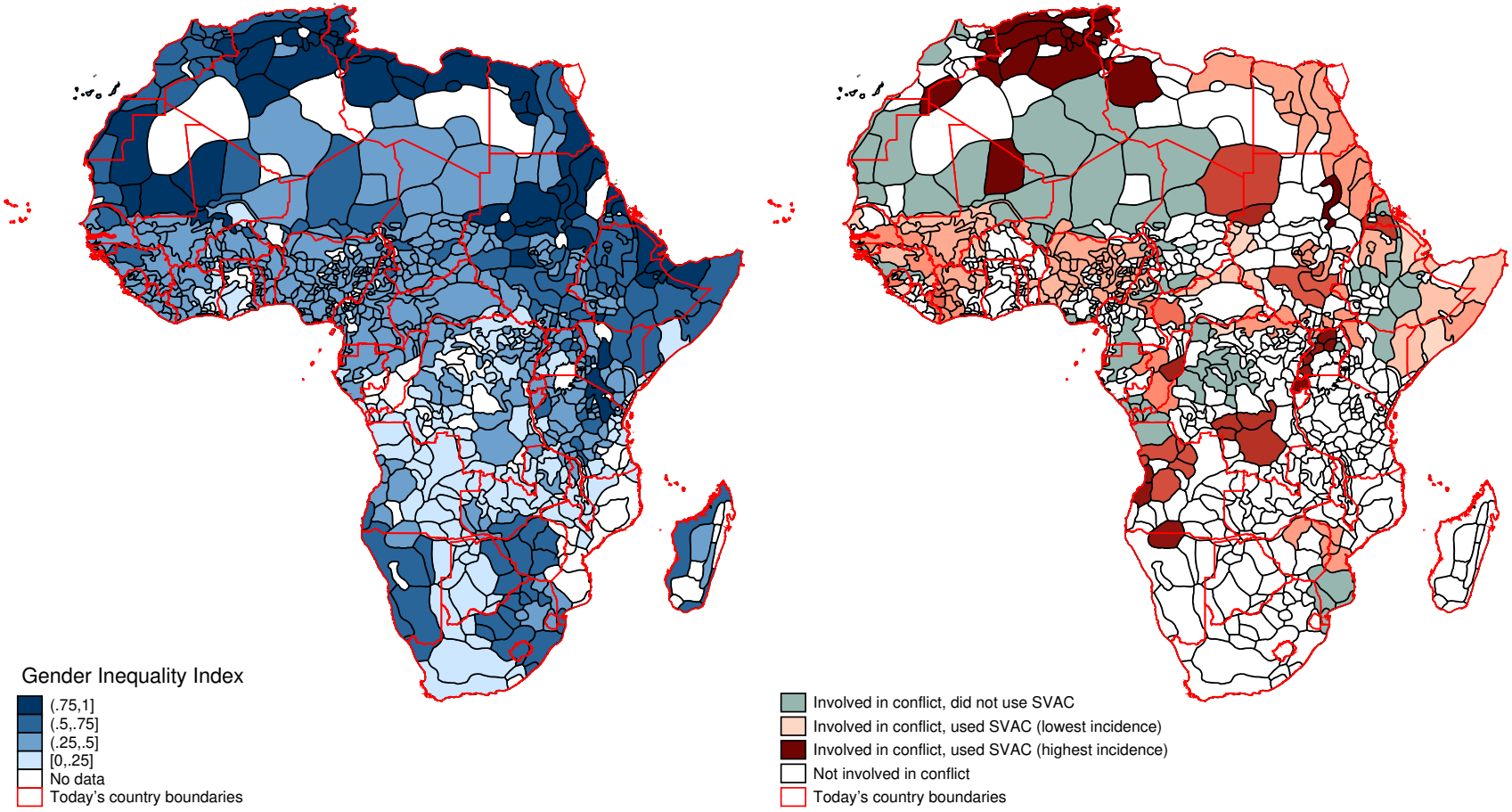
NOTES: Correlation between the eGII and reporting that "women are often/always treated unequally by the police" in Africa (32 countries) at the Murdock ethnicity level. Correlation coefficient: 0.13***
Sources: Ethnographic Atlas and Afrobarometer.

FIGURE 13: *Correlation between the eGII and reporting that "women are often/always treated unequally by the employer"*



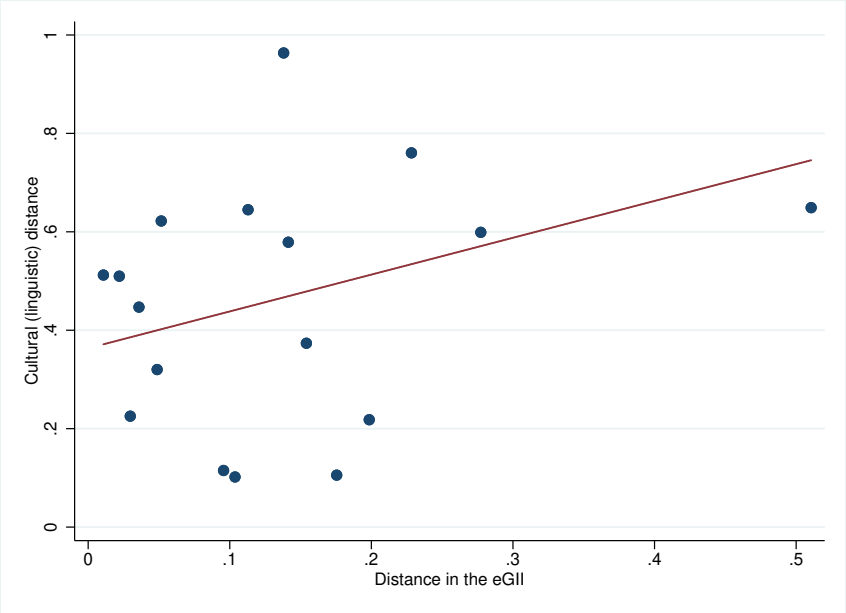
NOTES: Correlation between the eGII and reporting that "women are often/always treated unequally by the employer" in Africa (32 countries) at the Murdock ethnicity level. Correlation coefficient: 0.14***
Sources: Ethnographic Atlas and Afrobarometer.

FIGURE 14: *Distribution of the eGII across Africa and incidence of sexual violence in armed conflict*



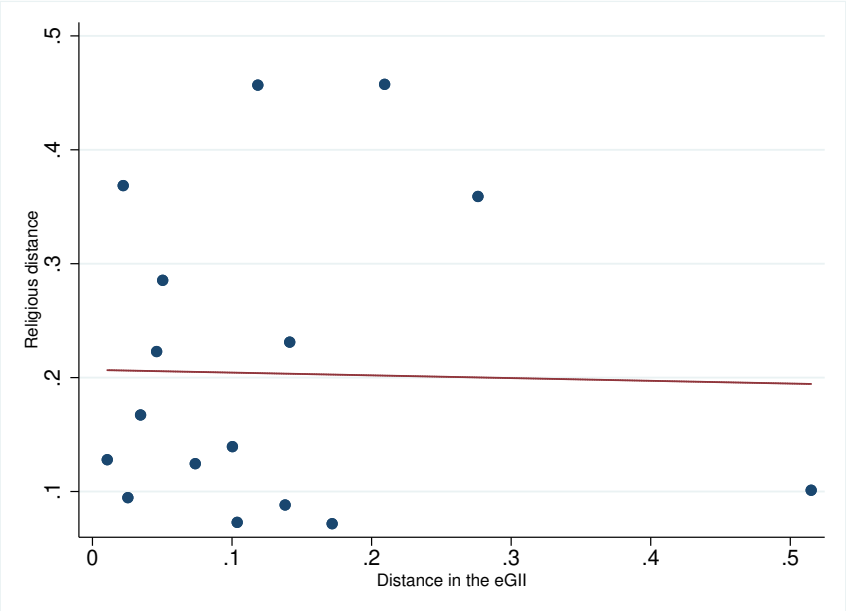
NOTES: Left: Distribution of the eGII across Murdock's ethnicities in Africa and contemporary country borders. Right: Total incidence of the use of sexual violence in armed conflict by Murdock ethnicities from 1989 to 2009, measured through an index ranging between 0 and 1.

FIGURE 15: *Correlation between cultural distance in gender norms and overall cultural (linguistic) distance*



NOTES: Correlation between the absolute distance in gender norms between the combatants and their cultural (linguistic) distance for the sample of ethnicities involved in inter-ethnic conflict. Correlation coefficient: 0.25***. Sources: Murdock Ethnographic Atlas and Ethnologue.

FIGURE 16: *Correlation between cultural distance in gender norms and religious distance*



NOTES: Correlation between the absolute distance in gender norms between the combatants and their religious distance for the sample of ethnicities involved in inter-ethnic conflict. Correlation coefficient: -0.02. Sources: Murdock Ethnographic Atlas and EPR-ED dataset.

TABLE 1: *Dataset extract illustrating the dyadic structure*

Country	**Year**	**Conflict ID**	**Perpetrator**	**Victim**	**Perpetrator's ethnicity**	**Victim's ethnicity**	**SVAC**
Chad	1994	288	Government of Chad	CSNPD	Zaghawa, Bideyat	Sara	2
Chad	1994	288	CSNPD	Government of Chad	Sara	Zaghawa, Bideyat	0

TABLE 2: *Perpetrator's ethnic characteristics and sexual violence in armed conflict*

	Dependent variable: sexual violence (0-3)					
	(1)	(2)	(3)	(4)	(5)	(6)
Lineage, Residence and Family Arrangements						
Matrilineal (weighted)	-0.78***	-0.85***	-0.66**			
	(0.219)	(0.283)	(0.326)			
Matrilineal				-0.77***	-0.84***	-0.66**
				(0.218)	(0.276)	(0.315)
Adjusted R^2	0.292	0.363	0.367	0.292	0.364	0.368
Patrilocal (weighted)	0.65***	0.70***	0.60**			
	(0.201)	(0.249)	(0.284)			
Patrilocal				0.64***	0.69***	0.59**
				(0.205)	(0.248)	(0.286)
Adjusted R^2	0.291	0.362	0.363	0.291	0.361	0.362
Stem (weighted)	-0.42	-0.47*	-0.40			
	(0.340)	(0.283)	(0.521)			
Stem				-0.50	-0.52*	-0.28
				(0.334)	(0.301)	(0.533)
Adjusted R^2	0.278	0.347	0.347	0.282	0.350	0.351
Mean dep. var.	0.62	0.62	0.62	0.62	0.62	0.62
Conflict fixed effect	✓	✓	✓	✓	✓	✓
Year fixed effect	✓	✓	✓	✓	✓	✓
Conflict-specific time trend		✓	✓		✓	✓
Victim's characteristic			✓			✓
Observations	900	900	893	900	900	893
Clusters	128	128	127	128	128	127

NOTES: OLS coefficient estimates of Equation 1. The dependent variable is an index ranging between 0 and 3 that captures the intensity of sexual violence. Explanatory variables are perpetrator-specific ethnic characteristics—either unweighted or weighted by the ethnic group's land area—capturing lineage systems (matrilineal), residence patterns (virilocal) and family arrangements (stem). All explanatory variables are normalized and range between 0 and 1. Standard errors are clustered at the perpetrator's level. *** (**) (*) indicate significance at the 1% (5%) (10%) level.

TABLE 3: *Perpetrator's ethnic characteristics and sexual violence in armed conflict*

	Dependent variable: sexual violence (0-3)					
	(1)	(2)	(3)	(4)	(5)	(6)
Subsistence Activities (I)						
Gathering (weighted)	1.70	1.40	1.55			
	(1.431)	(1.522)	(1.560)			
Gathering				0.71	0.27	0.82
				(1.823)	(2.000)	(2.246)
Adjusted R^2	0.274	0.340	0.341	0.271	0.338	0.340
Hunting (weighted)	1.31	1.38	2.77			
	(2.212)	(2.215)	(2.703)			
Hunting				2.50	2.46	3.40
				(2.275)	(2.294)	(2.764)
Adjusted R^2	0.272	0.340	0.343	0.276	0.343	0.345
Agriculture (weighted)	-1.38**	-1.34**	-1.47*			
	(0.569)	(0.605)	(0.877)			
Agriculture				-1.47***	-1.46**	-1.43*
				(0.545)	(0.578)	(0.786)
Adjusted R^2	0.295	0.362	0.363	0.300	0.367	0.368
Mean dep. var.	0.62	0.62	0.62	0.62	0.62	0.62
Conflict fixed effect	✓	✓	✓	✓	✓	✓
Year fixed effect	✓	✓	✓	✓	✓	✓
Conflict-specific time trend		✓	✓		✓	✓
Victim's characteristic			✓			✓
Observations	900	900	893	900	900	893
Clusters	128	128	127	128	128	127

NOTES: OLS coefficient estimates of Equation 1. The dependent variable is an index ranging between 0 and 3 that captures the intensity of sexual violence. Explanatory variables are perpetrator-specific ethnic characteristics—either unweighted or weighted by the ethnic group's land area—capturing dependence on different subsistence activities (gathering, hunting, agriculture). All explanatory variables are normalized and range between 0 and 1. Standard errors are clustered at the perpetrator's level. *** (**) (*) indicate significance at the 1% (5%) (10%) level.

TABLE 4: *Perpetrator's ethnic characteristics and sexual violence in armed conflict*

	Dependent variable: sexual violence (0-3)					
	(1)	(2)	(3)	(4)	(5)	(6)
Subsistence Activities (II)						
Plough (weighted)	0.15	0.07	-0.05			
	(0.606)	(0.701)	(0.701)			
Plough				0.31	-0.23	0.09
				(0.689)	(0.506)	(0.781)
Adjusted R^2	0.274	0.340	0.341	0.271	0.338	0.340
Husbandry (weighted)	1.50***	1.56***	1.56*			
	(0.567)	(0.578)	(0.812)			
Husbandry				1.65***	1.70***	1.59**
				(0.506)	(0.523)	(0.713)
Adjusted R^2	0.296	0.366	0.367	0.304	0.375	0.375
Pastoralism (weighted)	1.50***	1.56***	1.57*			
	(0.562)	(0.575)	(0.810)			
Pastoralism				1.67***	1.73***	1.62**
				(0.504)	(0.522)	(0.713)
Adjusted R^2	0.297	0.367	0.367	0.305	0.376	0.376
Mean dep. var.	0.62	0.62	0.62	0.62	0.62	0.62
Conflict fixed effect	✓	✓	✓	✓	✓	✓
Year fixed effect	✓	✓	✓	✓	✓	✓
Conflict-specific time trend		✓	✓		✓	✓
Victim's characteristic			✓			✓
Observations	900	900	893	900	900	893
Clusters	128	128	127	128	128	127

NOTES: OLS coefficient estimates of Equation 1. The dependent variable is an index ranging between 0 and 3 that captures the intensity of sexual violence. Explanatory variables are perpetrator-specific ethnic characteristics—either unweighted or weighted by the ethnic group's land area—capturing the use of the plough, and dependence on different subsistence activities (animal husbandry and pastoralism). The specification with "use of the plough" as an explanatory variable controls for dependence on agriculture. All explanatory variables are normalized and range between 0 and 1. Standard errors are clustered at the perpetrator's level. *** (**) (*) indicate significance at the 1% (5%) (10%) level.

TABLE 5: *Perpetrator's ethnic characteristics and sexual violence in armed conflict*

	Dependent variable: sexual violence (0-3)					
	(1)	(2)	(3)	(4)	(5)	(6)
The Slave Trade						
Transatlantic Slave Trade	-1.12**	-0.86*	-1.56***			
	(0.434)	(0.442)	(0.372)			
Indian Ocean Slave Trade				0.74***	0.74**	1.66
				(0.258)	(0.292)	(1.247)
Mean dep. var.	0.62	0.62	0.62	0.62	0.62	0.62
Conflict fixed effect	✓	✓	✓	✓	✓	✓
Year fixed effect	✓	✓	✓	✓	✓	✓
Conflict-specific time trend		✓	✓		✓	✓
Victim's characteristic			✓			✓
Observations	900	900	893	900	900	893
Clusters	128	128	127	128	128	127
Adjusted R^2	0.277	0.341	0.343	0.273	0.340	0.277

NOTES: OLS coefficient estimates of Equation 1. The dependent variable is an index ranging between 0 and 3 that captures the intensity of sexual violence. Explanatory variables are perpetrator-specific exposure to the Atlantic and the Indian slave trade, respectively. Both variables are constructed as: ln(1+Number of slaves/Ethnic group's land area) as in (Nunn and Wantchekon [2011]). All explanatory variables are normalized and range between 0 and 1. Standard errors are clustered at the perpetrator's level. *** (**) (*) indicate significance at the 1% (5%) (10%) level.

TABLE 6: *eGII: PCA loadings*

Variables	*Loading*
Gender Equal Traits	
Matrilineal	-0.26
Dependence on agriculture	-0.27
Gender Unequal Traits	
Virilocal	0.30
Dependence on pastoralism	0.55
Use of the plough	0.29
Dependence on animal husbandry	0.55
Ambiguous Traits	
Stem	0.01
Dependence on gathering	-0.15
Dependence on hunting	-0.26
Kaiser-Meyer-Olkin's measure of sampling adequacy	0.58

NOTES: Loadings from the principal component analysis on the eGII.

TABLE 7: *Gender Inequality Index and sexual violence in armed conflict*

	Dependent variable: sexual violence (0-3)					
	(1)	(2)	(3)	(4)	(5)	(6)
Ethnic Gender Inequality Index						
eGII (weighted)	1.83***	1.87***	1.55**			
	(0.505)	(0.564)	(0.715)			
eGII (unweighted)				1.84***	1.90***	1.54**
				(0.457)	(0.507)	(0.635)
Adjusted R^2	0.307	0.377	0.376	0.311	0.382	0.382
restricted eGII (weighted)	1.40***	1.43***	1.45*			
	(0.502)	(0.525)	(0.755)			
restricted eGII (unweighted)				1.54***	1.57***	1.46**
				(0.455)	(0.480)	(0.658)
Adjusted R^2	0.299	0.368	0.369	0.307	0.376	0.377
Mean dep. var.	0.62	0.62	0.62	0.62	0.62	0.62
Conflict fixed effect	✓	✓	✓	✓	✓	✓
Year fixed effect	✓	✓	✓	✓	✓	✓
Conflict-specific time trend		✓	✓		✓	✓
Victim's characteristic			✓			✓
Observations	900	900	893	900	900	893
Clusters	128	128	127	128	128	127

NOTES: OLS coefficient estimates of Equation 1. The dependent variable is an index ranging between 0 and 3 that captures the intensity of sexual violence. The explanatory variables are the perpetrator's eGII (weighted by the ethnic group land area and unweighted) and the perpetrator's restricted version of the eGII (weighted by the ethnic group land area and unweighted). All explanatory variables are normalized and range between 0 and 1. Standard errors are clustered at the perpetrator's level. *** (**) (*) indicate significance at the 1% (5%) (10%) level.

TABLE 8: *Cultural distance in gender norms and sexual violence in armed conflict (weighted eGII)*

	Dependent variable: sexual violence (0-3)				
	(1)	(2)	(3)	(4)	(5)
Absolute distance ($\lvert eGII_p - eGII_v \rvert$)	1.53***				
	(0.518)				
Perpetrator's eGII		0.58	2.05**	1.13	
		(0.629)	(0.957)	(0.997)	
Perpetrator *more* unequal		1.44*		1.51*	1.53***
		(0.814)		(0.811)	(0.503)
Perpetrator *less* unequal			0.20	0.64	1.56
			(0.919)	(0.888)	(1.214)
Conflict fixed effect	✓	✓	✓	✓	✓
Year fixed effect	✓	✓	✓	✓	✓
Conflict-specific time trends	✓	✓	✓	✓	✓
Perpetrator fixed effect	✓				✓
Mean dep. var.	0.62	0.62	0.62	0.62	0.62
Observations	623	643	643	643	623
Clusters	76	76	76	76	76
Adjusted R^2	0.597	0.379	0.374	0.379	0.596

NOTES: OLS coefficient estimates of Equations 3 and 4. The sample is restricted to include inter-ethnic conflicts only. The dependent variable is an index ranging between 0 and 3 that captures the intensity of sexual violence. The explanatory variables are the following: the absolute distance in the eGII between perpetrator and victim; the perpetrator's eGII (weighted by the ethnic group land area); the absolute distance in the eGII between perpetrator and victim when the perpetrator is more gender unequal than the victim; the absolute distance in the eGII between perpetrator and victim when the perpetrator is less gender unequal than the victim. Standard errors are clustered at the dyad level. *** (**) (*) indicate significance at the 1% (5%) (10%) level.

TABLE 9: *Cultural distance in gender norms and general violence: perpetrator's killings (weighted eGII)*

	Dependent variable: victim's fatalities (0-3)				
	(1)	(2)	(3)	(4)	(5)
Absolute distance ($\lvert eGII_p - eGII_v \rvert$)	-0.18				
	(0.856)				
Perpetrator's eGII		-0.03	-1.40	-0.63	
		(0.720)	(0.991)	(1.136)	
Perpetrator *more* unequal		-1.17		-1.24	-0.06
		(0.807)		(0.805)	(0.976)
Perpetrator *less* unequal			-0.35	-0.71	-0.58
			(1.011)	(0.982)	(1.111)
Conflict fixed effect	✓	✓	✓	✓	✓
Year fixed effect	✓	✓	✓	✓	✓
Conflict-specific time trends	✓	✓	✓	✓	✓
Perpetrator fixed effect	✓				✓
Mean dep. var.	1.07	1.07	1.07	1.07	1.07
Observations	623	643	643	643	623
Clusters	76	76	76	76	76
Adjusted R^2	0.317	0.266	0.263	0.265	0.316

NOTES: OLS coefficient estimates of Equations 3 and 4. The sample is restricted to include inter-ethnic conflicts only. The dependent variable is an index ranging between 0 and 3 that captures the intensity of deaths inflicted by the perpetrator on the victim, coded like the sexual violence variable (0: no killings; 1: between 1 and 24; 2 between 25 and 999; 3: equal to or larger than 1000). The explanatory variables are the following: the absolute distance in the eGII between perpetrator and victim; the perpetrator's eGII; the absolute distance in the eGII between perpetrator and victim when the perpetrator is more gender unequal than the victim; the absolute distance in the eGII between perpetrator and victim when the perpetrator is less gender unequal than the victim. Standard errors are clustered at the dyad level. *** (**) (*) indicate significance at the 1% (5%) (10%) level.

TABLE 10: *Cultural distance in gender norms, linguistic distance, and sexual violence in armed conflict*

	Dependent variable: sexual violence (0-3)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Distance in gender norms ($\lvert eGII_p - eGII_v \rvert$)	1.53***	1.52***	1.54***					
	(0.518)	(0.513)	(0.516)					
Perpetrator *more* gender unequal				1.53***	1.51***	1.53***		
				(0.504)	(0.502)	(0.500)		
Perpetrator *less* gender unequal				1.56	1.54	1.56		
				(1.214)	(1.205)	(1.209)		
Distance in other cultural traits (residuals)		-0.02			-0.02		-0.08	
		(0.050)			(0.049)		(0.107)	
Linguistic distance			-0.02			-0.02		-0.01
			(0.050)			(0.049)		(0.072)
Conflict fixed effect	✓	✓	✓	✓	✓	✓	✓	✓
Year fixed effect	✓	✓	✓	✓	✓	✓	✓	✓
Conflict-specific time trends	✓	✓	✓	✓	✓	✓	✓	✓
Perpetrator fixed effect	✓	✓	✓	✓	✓	✓	✓	✓
Mean dep. var.	0.62	0.62	0.62	0.62	0.62	0.62	0.62	0.62
Observations	623	623	623	623	623	623	623	623
Clusters	76	76	76	76	76	76	76	76
Adjusted R^2	0.597	0.596	0.596	0.596	0.596	0.596	0.592	0.592

NOTES: OLS coefficient estimates of Equations 3 and 4 with controls for distance in other cultural traits or linguistic distance. The sample is restricted to include inter-ethnic conflicts only. The dependent variable is an index ranging between 0 and 3 that captures the intensity of sexual violence. The explanatory variables are the following: the absolute distance in the eGII between perpetrator and victim; the absolute distance in the eGII between perpetrator and victim when the perpetrator is more gender unequal than the victim; the absolute distance in the eGII between perpetrator and victim when the perpetrator is less gender unequal than the victim; residuals of regressing linguistic distance on distance in gender norms; linguistic distance between perpetrator and victim. Standard errors are clustered at the dyad level. *** (**) (*) indicate significance at the 1% (5%) (10%) level.

TABLE 11: *Cultural distance in gender norms, religious distance, and sexual violence in armed conflict*

	Dependent variable: sexual violence (0-3)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Distance in gender norms ($\lvert eGII_p - eGII_v \rvert$)	1.53***	1.65***	1.65***					
	(0.518)	(0.552)	(0.556)					
Perpetrator *more* gender unequal				1.53***	1.96***	1.96***		
				(0.504)	(0.465)	(0.470)		
Perpetrator *less* gender unequal				1.56	1.14	1.14		
				(1.214)	(1.114)	(1.114)		
Distance in other cultural traits (residuals)		-0.06			0.03		-0.44**	
		(0.182)			(0.177)		(0.212)	
Religious distance			-0.06			0.03		-0.45**
			(0.182)			(0.177)		(0.209)
Conflict fixed effect	✓	✓	✓	✓	✓	✓	✓	✓
Year fixed effect	✓	✓	✓	✓	✓	✓	✓	✓
Conflict-specific time trends	✓	✓	✓	✓	✓	✓	✓	✓
Perpetrator fixed effect	✓	✓	✓	✓	✓	✓	✓	✓
Mean dep. var.	0.62	0.62	0.62	0.62	0.62	0.62	0.62	0.62
Observations	623	590	590	623	590	590	590	590
Clusters	76	72	72	76	72	72	72	72
Adjusted R^2	0.597	0.560	0.560	0.596	0.559	0.559	0.556	0.556

NOTES: OLS coefficient estimates of Equations 3 and 4 with controls for distance in other cultural traits or religious distance. The sample is restricted to include inter-ethnic conflicts only. The dependent variable is an index ranging between 0 and 3 that captures the intensity of sexual violence. The explanatory variables are the following: the absolute distance in the eGII between perpetrator and victim; the absolute distance in the eGII between perpetrator and victim when the perpetrator is more gender unequal than the victim; the absolute distance in the eGII between perpetrator and victim when the perpetrator is less gender unequal than the victim; residuals of regressing religious distance on distance in gender norms; religious distance between perpetrator and victim. Standard errors are clustered at the dyad level. *** (**) (*) indicate significance at the 1% (5%) (10%) level.

Appendix A

A-1 Data Sources and Dataset Construction

We hereby present in detail the data sources used for the analysis and the procedure adopted to merge the various datasets.

A-11. Sexual Violence in Armed Conflict

The source of our dependent variable is the Sexual Violence in Armed Conflict Dataset (SVAC) (Cohen and Nordås [2014]). The SVAC dataset includes information on all conflicts between 1989 and 2009, as defined by the UCDP/PRIO Armed Conflict Database: any "contested incompatibility that concerns government and/or territory where the use of armed force between two parties, of which at least one is the government of a state, results in at least 25 battle-related deaths" (Gleditsch *et al.* [2002]). The SVAC dataset provides information on war-related sexual violence perpetrated by three types of armed-conflict actors: government/state military, pro-government militias, and rebel/insurgent forces. In total, the dataset covers 129 active conflicts and 625 armed actors involved in them. Adhering to the International Criminal Court's rationale, SVAC defines war-related sexual violence as including the following acts: rape, sexual slavery, forced prostitution, forced pregnancy, forced sterilization, and forced abortion (International Criminal Court [2000]). In addition, following Wood [2009], sexual mutilation and sexual torture are also included.

SVAC draws upon annual reports from three sources (Amnesty International, Human Rights Watch, and the US State Department) to construct a measure of prevalence of sexual violence at the conflict-actor-year level. The resulting variable is an index ranging between 0-3 that reflects the magnitude of the phenomenon. More specifically, it takes the value 3 if, in a given year of conflict, an actor perpetrated acts of massive, innumerable, or systematic sexual violence according to the aforementioned sources and, furthermore, if reported incidents or victims of sexual violence exceeded 1,000; 2 if sexual violence was described as widespread and common, and reports of victims or incidents ranged between 25 and 999; 1 if reported victims and incidents were below 25 and the occurrence of sexual violence was only isolated; 0 if no sexual violence was mentioned in a given year in relation to a specific conflict.[22]

We exploit the dyadic version of the SVAC dataset, i.e., the GEO-SVAC dataset (Bahgat *et al.* [2016]). GEO-SVAC uses as its starting point the UCDP GED dataset (Sundberg and Melander [2013]; Croicu and Sundberg [2017]) and enriches it with the information on sexual violence provided by SVAC for *state-based* conflicts between government/state military and rebel/insurgent forces between 1989 and 2009.[23]

[22]For further details on the methodology of data collection and coding refer to Cohen and Nordås [2014].

[23]Figure A-1 illustrates the relationship between GEO-SVAC, UCDP GED, and SVAC. Conflicts covered by GEO-SVAC are only a subset of those included in the original UCDP GED dataset, which includes also *non-state* conflicts and episodes of *one-sided* violence. Moreover, GEO-SVAC does not cover SVAC conflict events involving pro-government militias. As a result, GEO-SVAC includes information on 106

The unit of observation in GEO-SVAC is a geo-located state-based conflict event. Since the variation in our variable of interest (i.e., sexual violence prevalence) occurs at the actor-conflict-year level—and not at a geo-located event level—we maintain actor-conflict-year as the unit of observation in the analysis. For our purposes, however, GEO-SVAC offers an important advantage. In addition to the identity of the perpetrator of sexual violence, it codes the identity of the other actor involved in the conflict. In other words, the dataset is *dyadic*, i.e. it includes the identity of side A, which is always a government (corresponding "government/state military" in the SVAC coding) and side B, a rebel or opposing government (corresponding to "rebel/insurgent forces" in SVAC). It furthermore reports the intensity of sexual violence perpetrated by both side A and side B in a specific year of conflict.

As illustrated in Figure A-2, we restrict the GEO-SVAC sample in two ways. First, we focus on 45 conflicts fought in the African continent.[24] Second, we restrict our analysis to 33 *ethnic* civil conflicts defined as "armed conflicts between the government of a state and one or more internal opposition group(s) that cause at least 25 battle-related deaths within a year and in which armed groups (i) explicitly pursue ethno-nationalist aims, motivations, and interests; and (ii) recruit fighters and forge alliances on the basis of ethnic affiliation" (Gleditsch *et al.* [2002], Cederman *et al.* [2012]).[25] We categorize conflict-years as ethnic relevant based on Wimmer *et al.*'s [2009] definition.[26] In addition, we include in the sample three additional conflicts that, according to the sources we consulted, qualify as ethnic.[27] Our results are robust to dropping the latter three conflicts from the sample.

A-12. Conflict Actors' Ethnic Identity

Next, we assign to each actor (i.e., to both side A and side B in GEO SVAC) an ethnic identity. To achieve this, we exploit the rich information provided by the Ethnic Power Relations (EPR) Dataset Family (Vogt *et al.* [2015]), where an ethnic group is "an identity group that defines itself or is defined by others along linguistic, religious, or racial characteristics". The EPR dataset family provides information, *inter alia*, on ethnic groups' involvement in civil war as part of a rebel organization. We are therefore able to assign

state-based conflicts around the world involving the following actors: *government/state military* and *rebel/insurgent forces*. Finally, GEO-SVAC includes only active years of conflict, whereas SVAC provides information also on interim and post-conflict years.

[24] African conflicts constitute 42% of conflicts in GEO-SVAC, which includes a total of 106 conflicts.

[25] "[...] we conducted new research and coded each conflict for whether rebel organizations pursued ethnonationalist aims and recruited along ethnic lines. We also coded whether rebels aimed at establishing a new independent state. We distinguish between ethnic and nonethnic conflicts using the aims of the armed organization and their recruitment and alliance structures [...]. We identify as "ethnic" the aims of achieving ethnonational self-determination, a more favorable ethnic balance-of-power in government, ethnoregional autonomy, the end of ethnic and racial discrimination, language and other cultural rights, and so forth. In ethnic wars, armed organizations also recruit fighters predominantly among their leaders' ethnic group and forge alliances on the basis of ethnic similarity" (Wimmer *et al.*'s [2009]).

[26] In the case of 17 conflicts categorized as ethnic in Wimmer *et al.* [2009], we include additional conflict-years that were not recorded by Wimmer *et al.* [2009] but that were part of a conflict that was qualified as ethnic. Results are robust to excluding these conflict years.

[27] Government of Guinea-Bissau vs. Military Junta for the Consolidation of Democracy, Peace and Justice (1998-1999); Government of Eritrea vs. Government of Ethiopia (1998-2000); Government of Eritrea vs. EIJM-AS (1993-2003).

FIGURE A-1: *Relationship between UCDP-GED, SVAC and GEO-SVAC*

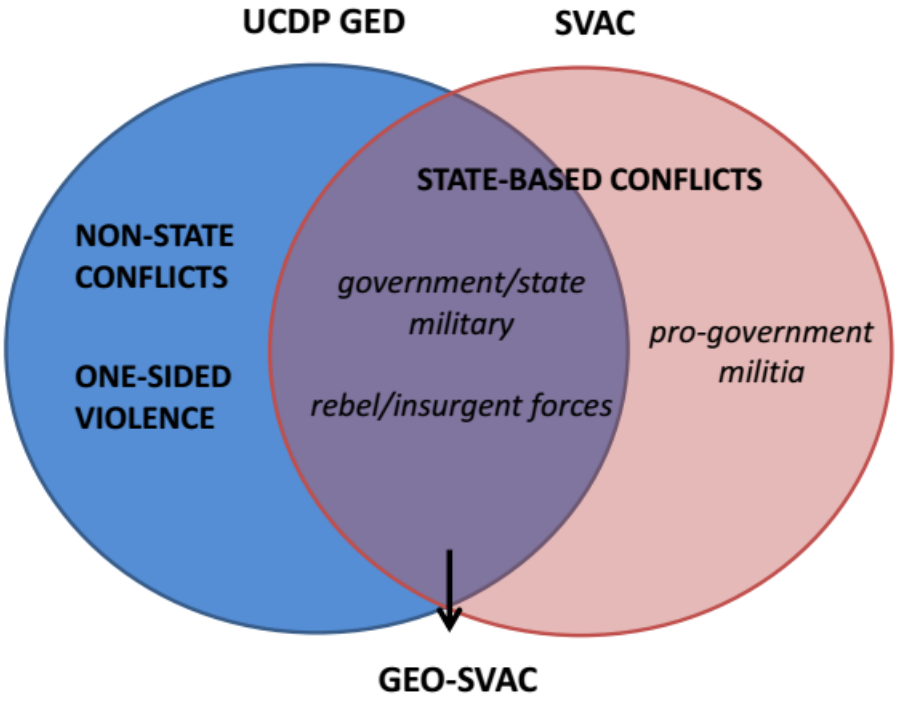
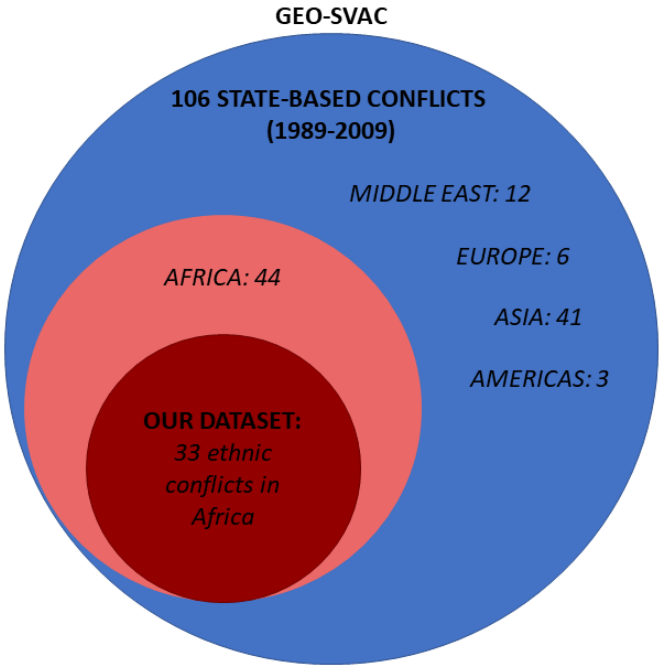


FIGURE A-2: *Relationship between GEO-SVAC and our dataset*



to the majority of sides B (i.e. rebel or insurgent forces as mentioned above) one or more EPR ethnic groups.[28] Using a variety of additional sources, we identify the ethnic identity of four remaining sides B involved in conflicts classified as ethnic.[29]

The EPR dataset family also provides information on ethnic groups' access to executive government power. When an ethnic group holds exclusive or almost exclusive power in the government, it is classified as either *Monopolist* or *Dominant*. When power is formally or informally shared by different groups, the latter are defined as *Senior* or *Junior Partners*, depending on their relative position in the government. Groups that do not detain any power are either defined as *Discriminated* or *Powerless*, depending on whether or not the central power pursues actions of active discrimination against them. The remaining categories refer to either *Self-Excluded*—i.e. controlling a particular territory in the state that they have declared independent—or *Irrelevant* groups. Finally, EPR also records instances of *State Collapse*.

Given the nature of conflicts included in our data—civil (state-based) ones where side A is always a government—we can assign to side A an ethnic identity. In instances where, in a certain year and during a conflict, central power is held exclusively by one ethnic group—defined by EPR either as *Dominant* or *Monopolist*—the matching is straightforward. Whenever more groups detain government power jointly, we always assign to side A the ethnic identity of the *Senior Partner*, and, in addition, of the *Junior Partner* only in cases where sources indicate direct involvement of that ethnic group in civil conflict. As a result, side A can be assigned to either one or more EPR ethnic groups.[30]

A-13. Conflict Actors' Ancestral Characteristics

Finally, we merge our dataset with the Ethnographic Atlas (EA), coded by Murdock [1967] and updated by Nunn and Wantchekon [2011]. The EA is arguably the most compelling source of ethnographic information for 1,265 societies around the world, collected at the end of the 19th century. For Africa, the EA provides detailed information on groups' socio-economic conditions, settlement patterns, and family arrangements prior to European contact.

We link the information provided by the EA to the dataset on conflict through the concordance data provided by Michalopoulos and Papaioannou [2016]. This data links 196 EPR groups to 593 ethnicities in Murdock using a variety of sources. We successfully merge 71 EPR groups to the EA through Michalopoulos and Papaioannou's [2016] concor-

[28]The sub-dataset—belonging to the EPR dataset family—that allows this merging is ACD2EPR (Wucherpfennig *et al.* [2012]). We classified the ethnic identity of 91 rebel groups with this procedure. To quality-check the validity of the merging, we simultaneously consulted the narratives in the EPR Atlas, accessible through the GrowUp database (https://growup.ethz.ch/). For 51 of these rebel groups, we confirmed the merging by consulting additional sources.

[29]These sides B belong to the three conflicts that Wimmer *et al.* [2009] do not classify as ethnic-relevant (the Military Junta for the Consolidation of Democracy, Peace, and Justice; EIJM-AS; Government of Ethiopia) and to a rebel group whose ethnic identity is missing in ACD2EPR (the AQIM in Algeria and Niger).

[30]We always conduct a quality-check on these merges by consulting the narratives in the EPR Atlas, accessible through the GrowUp database (https://growup.ethz.ch/) and, in the case of 13 (out of 28) governments, additional sources.

dance table. For 13 of the 15 EPR groups[31] that remain unmerged, we rely on a variety of sources and identify the Murdock groups of interest.[32] For two EPR ethnic groups (Americo-Liberians and Muslim Eritreans), it is impossible to identify a correspondence in the EA, and therefore they remain un-merged.

In some cases, this matching procedure results in a one-to-one mapping between EPR and the Ethnographic Atlas. For example, the ethnic group of the rebel force FLEC-FAC in Angola, the *Cabindan Mayombe*, is matched with the *Yombe* group in the EA. However, in other cases, a conflict actor is associated to multiple Murdock groups either because (i) side A, side B, or both are represented by multiple EPR groups, as described in the previous section, or (ii) an EPR group corresponds to multiple groups in the EA, or (iii) both. An example of the latter case is the following: the RFDG rebel group in Guinea is composed of members belonging to the EPR groups called *Malinke* and *Peul*. In turn, the Michalopoulos and Papaioannou's [2016] correspondence table matches *Malinke* to four Murdock groups (*Yalunka, Konyanke, Malinke,* and *Koranko*), and the *Peul* to three Murdock groups (*Foutadjalon, Sokoto, Liptako*). In these instances, we weight the ethnic characteristics of each EPR group by the size of the EA groups to which it corresponds. In the just-mentioned example, *Peul*'s dependence on pastoralism will be a weighted average between *Foutadjalon's, Sokoto's*, and *Liptako's* dependence on pastoralism, based on the three ethnic groups' size, proxied by the land area covered by their settlements. We provide estimates using both the weighted and the un-weighted version of the various ethnic characteristics, and show that our results are generally insensitive to this procedure.

A-14. Example

Figure A-3 summarizes the merging process described in this section for a conflict event that took place in 1989 in Liberia between the rebel group NPFL (National Patriotic Front of Liberia) and the government. The GrowUp platform[33] illustrated in Figure A-3 provides a summary of the ethnic power relations in Liberia in the year 1989. One ethnic group, the *Krahn (Guere)*, detains exclusive power in the government and is thus defined as *Dominant*. The remaining politically relevant groups (the *Americo-Liberians*, the *Gio* and the *Mano*) are all discriminated against. However, only the latter two are involved in a conflict, i.e., those marked by a star.[34] Consequently, the group *Krahn (Guere)* is assigned to side A (the government of Liberia), while side B corresponds to the *Gio* and *Mano* groups.

To confirm the validity of these matches, we consult the chapter on Liberia in the EPR Atlas (Girardin *et al.* [2015]). The following extract confirms the *Krahn* dominant position in the government:

[31]These EPR groups are: Afar, Americo-Liberians, Arabs, Arabs/Moors, Bembe, Christian Eritreans, Gio, Goula Isaas (Somali), Mandingo, Masalit, Muslim Eritreans, Somali, Sharawis, and Zaghawa

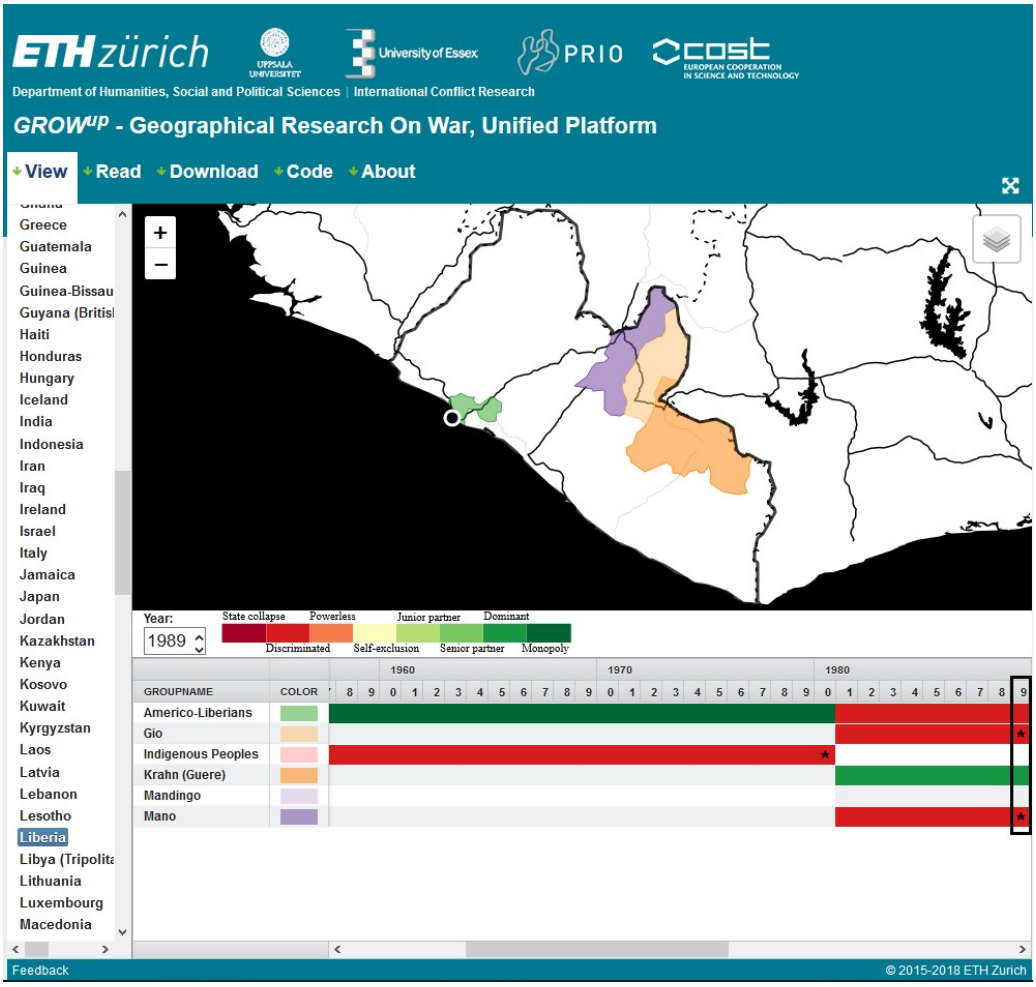
[32]Sources include the Joshua Project, the Ethnologue dataset, Wikipedia, and others. In some instances, we also exploit the fact that EPR provides the geo-location of the ethnic settlements to cross-validate the just mentioned sources.

[33]The interface depicted in Figure A-3 displays the information contained in the Ethnic Power Relations Dataset (Vogt *et al.* [2015]) and the UCDP conflict data (Croicu and Sundberg [2017]).

[34]This is the equivalent of the information contained in the ACD2EPR sub-dataset of the EPR family and in the Wimmer *et al.*'s [2009] classification of ethnic relevant conflict.

[...] Doe's coup brought an end to the Americo-Liberian dominance. [...] Doe's rule relied heavily on his own *Krahn* group, which occupied the state's key positions. They soon dominated political and military life in Liberia. Thus, the Krahn are coded as "dominant" during Doe's regime. There is also widespread discrimination and state violence against the *Gio* and *Mano* ethnic groups (where opposition against Doe was widespread) [...]. Thus, these groups are also coded as "discriminated".

FIGURE A-3: *Merging process example through the GrowUp platform*



NOTES: The figure displays the Grow-Up platform (Girardin *et al.* [2015]) with the Liberian example. On top, it displays the EPR groups' settlements, and at the bottom the power relations between the six ethnic groups from the 1960s to the 1980s. The year highlighted in black captures the war between the rebel forces represented by the Gio and Mano ethnicities and the government, represented by the Krahn (Guere) ethnic group.

Moreover, to confirm the ethnic nature of the NPFL rebel group, we rely on other narratives, such as Wikipedia, according to which "most NPFL fighters were originally drawn from the Gio and Mano ethnic groups of northern Liberia that were persecuted

under Doe's regime".[35]

Finally, the remaining step consists in associating the EPR groups with the EA. Straightforwardly, Michalopoulos and Papaioannou [2016] assign *Krahn* to *Kran* and *Mano* to *Ngere* in the Atlas. Gio is not included in the correspondence table. However, we retrieve the necessary information from Holsoe and Lauer [1976], according to whom "in Liberia, Gio persisted as the name for the Dan"[36], and link the EPR group *Gio* to *Dan* in the EA. As a final check, we also compare the EPR ethnic boundaries with the Ethnographic Atlas settlement map.

A-2 Linguistic Distance

We use Fearon's [2003] measure of linguistic distance, which is based on linguistic trees in the Ethnologue. For each language, the Ethnologue provides a classification starting with the language family (e.g. Afro Asiatic, Nilo-Saharan, Creole), followed by "nodes", i.e., the branching points of the linguistic tree, and ending with the language itself. We merge information on languages spoken by ethnic groups through the Ethnic Power Relations-Ethnic Dimensions (EPR-ED) dataset, and compute distances between each pair of languages based on the number of common nodes in the tree.

For example, the language spoken by the ethnic group Zaghawa is classified as follows: *Nilo-Saharan, Saharan, Eastern*. The language of Zaghawa's opponent, Sara, is classified as: *Nilo-Saharan, Satellite-Core, Satellites, Central Sudanic, West, Bongo-Bagirmi, Sara-Bagirmi, Sara, Sara Proper*. These two languages have only one node in common (Nilo-Saharan, i.e. the language family). Following Putterman and Weil [2010], we calculate the distance between language i and language j as follows:

$$d_{ij} = 1 - \left(\frac{\#\text{ of common nodes between } i \text{ and } j}{\frac{1}{2}(\#\text{ of nodes of language } i + \#\text{ of nodes of language } j)} \right)^{\lambda} \tag{5}$$

Languages originating from different families have no nodes in common, and their distance will be equal to 1. The parameter λ ranges between 0 and 1, and is used to attribute higher weight to earlier common nodes, as early separations in the language tree are likely to signify larger cultural divergence on average than later separations (see Fearon [2003]). As in Putterman and Weil [2010] and Fearon [2003], we assign to λ the value of 0.5.[37]

The EPR-ED dataset assigns to each EPR ethnic group up to three languages, which are the three largest language segments spoken by group members in descending order. It also attributes a relative size to each of these languages, which sums up to 1 and reflects the percentage of individuals within an ethnic group speaking a specific language. Given this, we exploit the relative size of languages as weights, and calculate the linguistic

[35] https://en.wikipedia.org/wiki/National_Patriotic_Front_of_Liberia

[36] Page 142 in Holsoe and Lauer [1976].

[37] In the above example, the linguistic distance between Zagawa and Sara is equal to:

$$d_{ij} = 1 - \left(\frac{1}{\frac{1}{2}(3+9)} \right)^{0.5} = 0.59 \tag{6}$$

distance for each perpetrator-victim pair of ethnic groups as follows:

$$LD_{pv} = \sum_{i=1}^{3} \sum_{j=1}^{3} (s_{pi} \times s_{vj} \times d_{ij}) \tag{7}$$

where p and v denote an ethnic group on the perpetrator's and on the victim's side, respectively, s_{pi} and s_{vj} denote the relative size of language i (j) in the ethnic group of the perpetrator (victim), and d_{ij} is the linguistic distance between language i and language j described above.

Since perpetrators and victims can be composed by multiple ethnic groups, the ultimate linguistic distance between two opposing actors in a conflict is given by the average distance between each perpetrator-victim ethnic-group pair:

$$LD_{PV} = \sum_{p=1}^{M} \sum_{v=1}^{N} (\frac{1}{M} \times \frac{1}{N} \times LD_{pv}) \tag{8}$$

where M denotes the number of ethnic groups fighting on the perpetrator's side, N the number of ethnic groups fighting on the victim's side, and LD_{pv} the linguistic distance of each ethnic group pair.

A-3 Religious Distance

We construct a measure of religious distance between ethnic belligerents exploiting information on ethnic groups' religion provided by the EPR-ED dataset. Similar to languages, EPR-ED codes up to three religions professed by each ethnic group, as well as their relative size (reflecting the percentage of individuals within an ethnic group professing a specific religion).

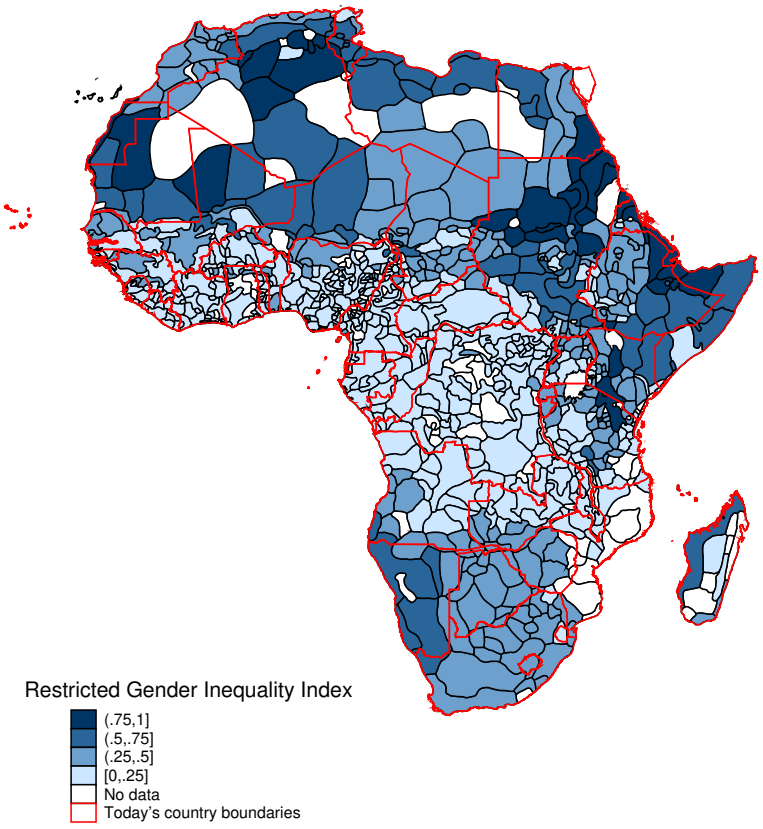
We construct a measure of religious distance analogous to the one for linguistic distance (see equations 5-8 in section A-2). To this end, we exploit EPR-ED classification of language segments. To continue the example of section A-2, the main religion of the ethnic group Zaghawa is Sunni Islam, classified as follows: *Abrahamic Religions, Islam, Sunni Islam.* The main religion of Zaghawa's opponent, Sara, is Protestantism, classified as *Abrahamic Religions, Christianity, Protestantism.* In this case, the two religions in the pair have one node in common, and their distance will be equal to 0.42.[38]

[38]Resulting from equation 5:

$$d_{ij} = 1 - \left(\frac{1}{\frac{1}{2}(3+3)} \right)^{0.5} = 0.42 \tag{9}$$

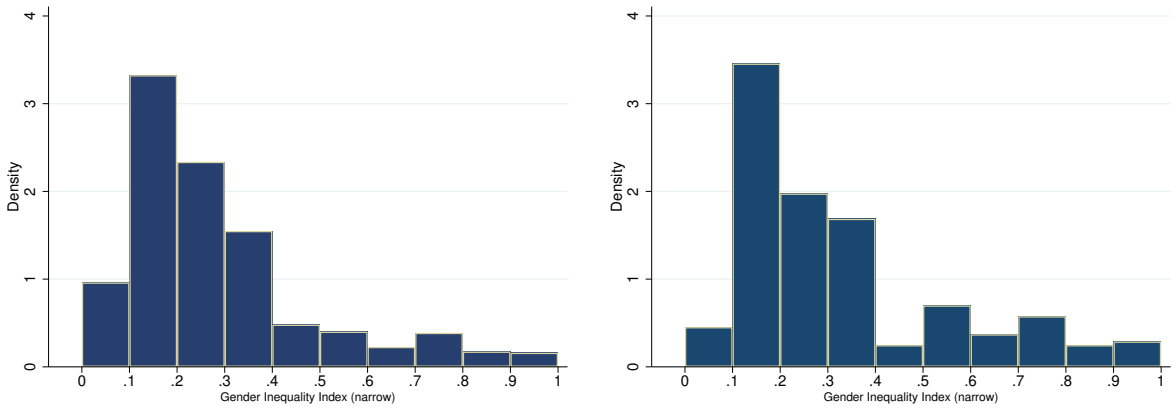
Appendix B: Additional Figures and Tables

FIGURE B-1: *Distribution of the restricted eGII across Africa*



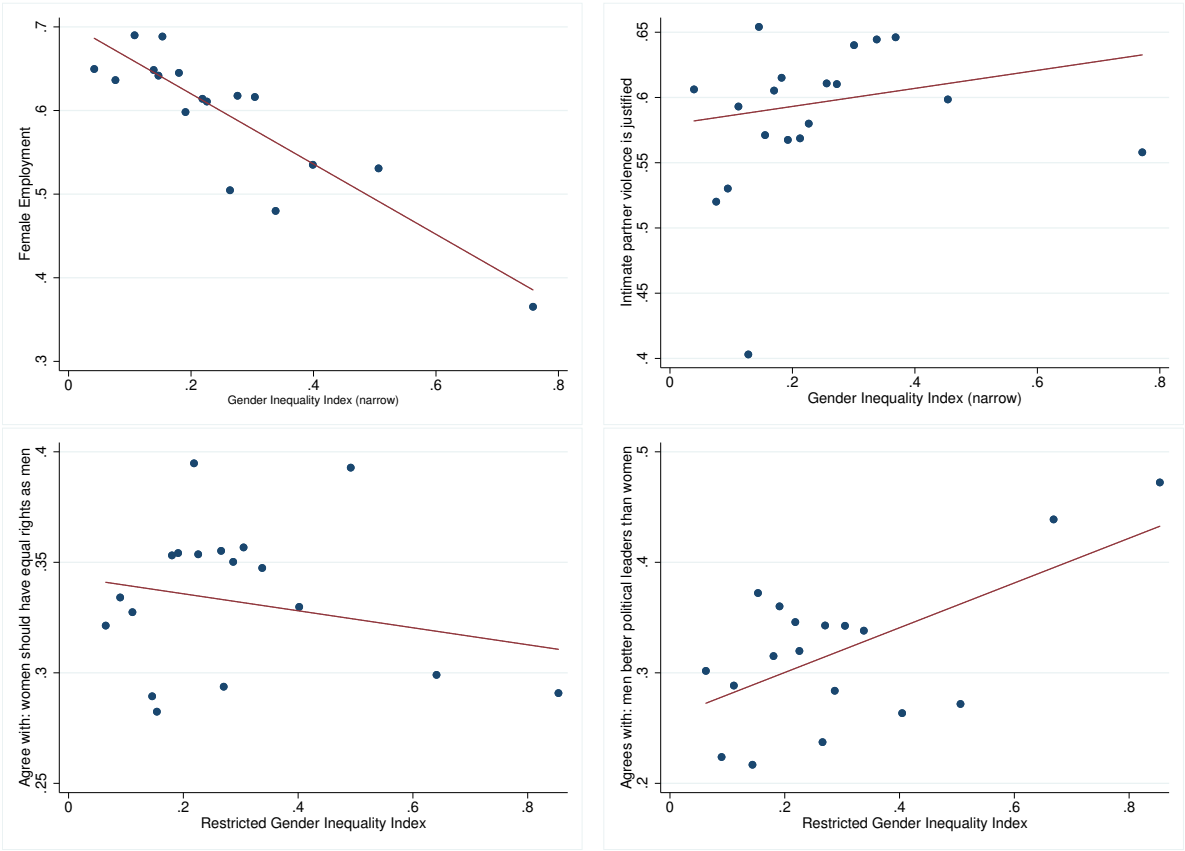
NOTES: Restricted Gender Inequality Index across Murdock's ethnicities in Africa and contemporary country borders.

FIGURE B-2: *Distribution of the restricted eGII*



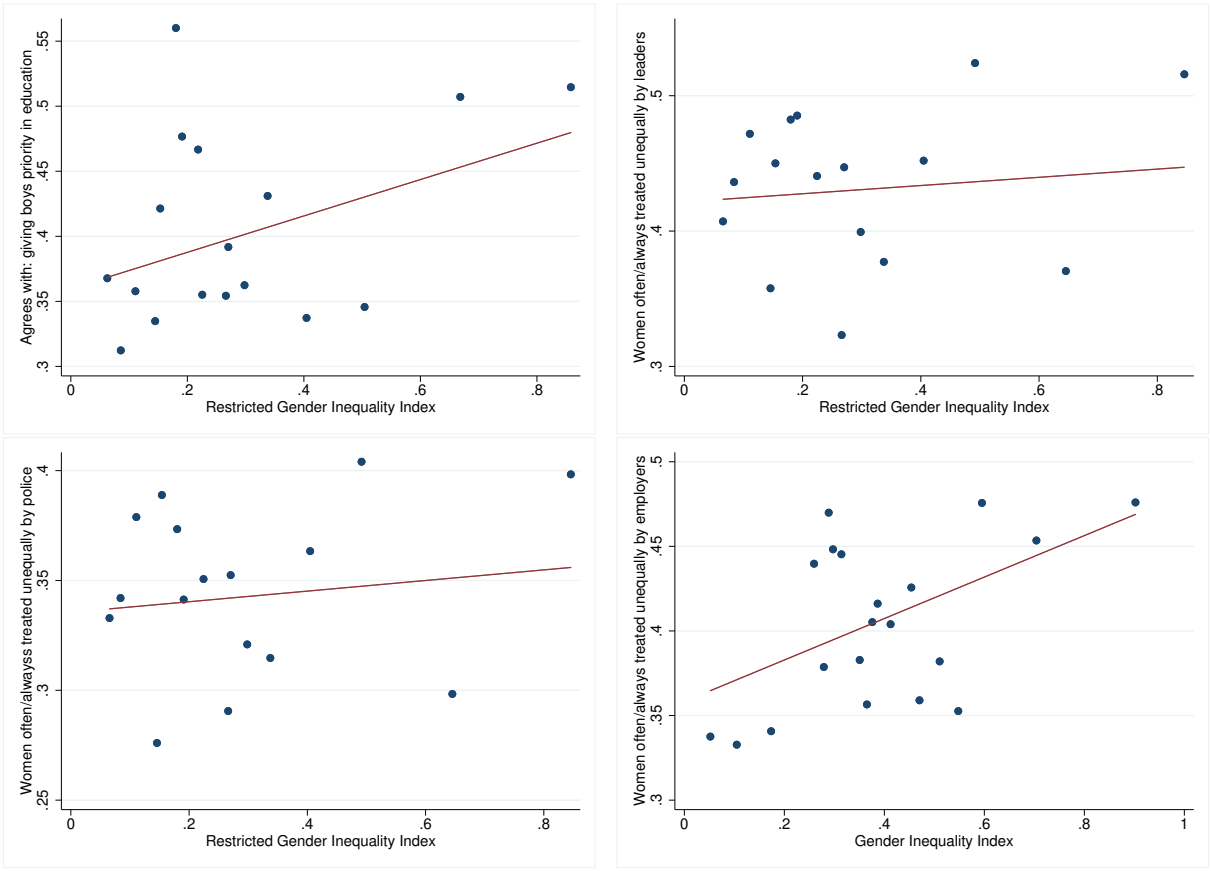
NOTES: Left: Distribution of the restricted eGII in Africa. Mean (standard deviation): 0.28 (0.20); right: Distribution of the restricted eGII in our sample. Mean (standard deviation): 0.32 (0.23)

FIGURE B-3: *Correlation between the restricted eGII and proxies for gender equality (ethnicity level)*



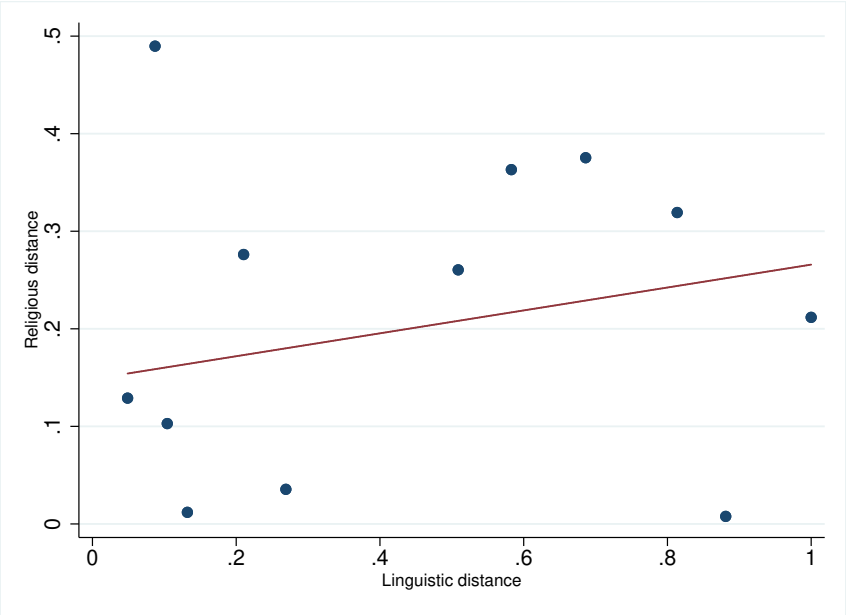
NOTES: Correlation between the restricted eGII and female employment (correlation coefficient: -0.36***) in sub-Saharan Africa (32 countries); the justification of intimate partner violence (correlation coefficient: 0.06) in sub-Saharan Africa (32 countries); agreeing with the statement "Women should have equal rights and receive the same treatment as men do" (correlation coefficient: -0.04) in Africa (35 countries); agreeing with the statement "Men make better political leaders than women" (correlation coefficient: 0.25***) in Africa (33 countries). Sources: Ethnographic Atlas, Demographic and Health (DHS) survey, and Afrobarometer.

FIGURE B-4: *Correlation between the restricted eGII and proxies for gender equality (ethnicity level)*



NOTES: Correlation between the restricted eGII and agreeing with the statement "If funds for schooling are limited, a boy should always receive an education in school before a girl" (correlation coefficient: 0.11**) in Africa (33 countries); reporting that "women are often/always treated unequally by leaders" (correlation coefficient: 0.03); reporting that "women are often/always treated unequally by the police" (correlation coefficient: 0.03); reporting that "women are often/always treated unequally by the employer" (correlation coefficient: 0.02). Sources: Ethnographic Atlas and Afrobarometer.

FIGURE B-5: *Correlation between linguistic distance and religious distance*



NOTES: Correlation between linguistic distance between the combatants and their religious distance for the sample of ethnicities involved in inter-ethnic conflict. Correlation coefficient: 0.23***. Sources: Ethnologue and EPR-ED dataset.

TABLE B-1: *Perpetrator's ethnic characteristics and sexual violence in armed conflict. Robustness tests*

	Dependent Variable: Sexual Violence (0-3)					
	Victim FE		Country FE		No time variation	
	Weighted	Unweighted	Weighted	Unweighted	Weighted	Unweighted
	(1)	(2)	(3)	(4)	(5)	(6)
Lineage, Residence and Family Arrangements						
Matrilineal	-0.33*	-0.35*	-0.84***	-0.82***	-0.58***	-0.55***
	(0.179)	(0.182)	(0.200)	(0.199)	(0.206)	(0.208)
Adjusted R^2	0.547	0.547	0.290	0.290	0.249	0.247
Virilocal	0.33*	0.31	0.83***	0.80***	0.51**	0.47**
	(0.182)	(0.192)	(0.186)	(0.185)	(0.213)	(0.216)
Adjusted R^2	0.547	0.546	0.299	0.296	0.244	0.242
Stem	0.06	0.42	-0.54**	-0.59**	-0.48	-0.71*
	(0.755)	(0.735)	(0.258)	(0.275)	(0.403)	(0.407)
Adjusted R^2	0.544	0.547	0.272	0.277	0.234	0.245
Conflict fixed effect	✓	✓			✓	✓
Country fixed effect			✓	✓		
Year fixed effect	✓	✓	✓	✓		
Conflict-specific time trend	✓	✓				
Country-specific time trend			✓	✓		
Victim fixed effect	✓	✓				
Mean dep. var.	0.62	0.62	0.62	0.62	0.57	0.57
Observations	880	880	900	900	266	266
Clusters	127	127	128	128	128	128

NOTES: OLS coefficient estimates of Equation 1. The dependent variable is an index ranging between 0 and 3 that captures the intensity of sexual violence. Explanatory variables are perpetrator-specific ethnic characteristics related to descent, residence patterns, family arrangements, and subsistence activities. All explanatory variables are normalized and range between 0 and 1. Observations are at the perpetrator-victim-conflict level. Columns (1), (3) and (5) report coefficients for covariates weighted by the size of the ethnic group, while columns (2), (4) and (6) report coefficients for unweighted covariates. Standard errors are clustered at the perpetrator's level. *** (**) (*) indicate significance at the 1% (5%) (10%) level.

TABLE B-2: *Perpetrator's ethnic characteristics and sexual violence in armed conflict. Robustness tests*

	Dependent Variable: Sexual Violence (0-3)					
	Victim FE		Country FE		No time variation	
	Weighted	Unweighted	Weighted	Unweighted	Weighted	Unweighted
	(1)	(2)	(3)	(4)	(5)	(6)
Subsistence Activities (I)						
Gathering	2.91	3.37	0.09	-0.71	1.52	0.35
	(1.790)	(2.301)	(1.536)	(1.841)	(1.318)	(1.489)
Adjusted R^2	0.548	0.547	0.259	0.259	0.230	0.226
Hunting	4.80*	4.66*	1.55	2.18	0.91	2.00
	(2.432)	(2.460)	(2.110)	(2.242)	(2.376)	(2.446)
Adjusted R^2	0.550	0.549	0.261	0.264	0.226	0.230
Agriculture	-1.48*	-1.30*	-1.02**	- 1.07**	-1.26**	-1.50***
	(0.845)	(0.715)	(0.483)	(0.472)	(0.588)	(0.572)
Adjusted R^2	0.551	0.550	0.279	0.280	0.251	0.264
Mean dep. var.	0.62	0.62	0.62	0.62	0.57	0.57
Conflict fixed effect	✓	✓			✓	✓
Country fixed effect			✓	✓		
Year fixed effect	✓	✓	✓	✓		
Conflict-specific time trend	✓	✓				
Country-specific time trend			✓	✓		
Victim fixed effect	✓	✓				
Observations	880	880	900	900	266	266
Clusters	127	127	128	128	128	128

NOTES: OLS coefficient estimates of Equation 1. The dependent variable is an index ranging between 0 and 3 that captures the intensity of sexual violence. Explanatory variables are perpetrator-specific ethnic characteristics related to descent, residence patterns, family arrangements, and subsistence activities. All explanatory variables are normalized and range between 0 and 1. Observations are at the perpetrator-victim-conflict level. Columns (1), (3) and (5) report coefficients for covariates weighted by the size of the ethnic group, while columns (2), (4) and (6) report coefficients for unweighted covariates. Standard errors are clustered at the perpetrator's level. *** (**) (*) indicate significance at the 1% (5%) (10%) level.

TABLE B-3: *Perpetrator's ethnic characteristics and sexual violence in armed conflict. Robustness tests*

	Dependent Variable: Sexual Violence (0-3)					
	Victim FE		Country FE		No time variation	
	Weighted	Unweighted	Weighted	Unweighted	Weighted	Unweighted
	(1)	(2)	(3)	(4)	(5)	(6)
Subsistence Activities (II)						
Plough	-0.78***	-0.74***	0.21	0.29	-0.47	-0.47
	(0.248)	(0.272)	(0.480)	(0.481)	(0.296)	(0.304)
Adjusted R^2	0.555	0.554	0.278	0.281	0.229	0.229
Husbandry	1.21	1.14*	1.09**	1.18***	1.24*	1.48**
	(0.778)	(0.654)	(0.458)	(0.440)	(0.657)	(0.603)
Adjusted R^2	0.549	0.549	0.282	0.286	0.248	0.263
Pastoralism	1.248	1.171*	1.080**	1.186***	1.24*	1.50**
	(0.779)	(0.654)	(0.456)	(0.439)	(0.648)	(0.602)
Adjusted R^2	0.549	0.550	0.282	0.286	0.249	0.264
Mean dep. var.	0.62	0.62	0.62	0.62	0.57	0.57
Conflict fixed effect	✓	✓			✓	✓
Country fixed effect			✓	✓		
Year fixed effect	✓	✓	✓	✓		
Conflict-specific time trend	✓	✓				
Country-specific time trend			✓	✓		
Victim fixed effect	✓	✓				
Observations	880	880	900	900	266	266
Clusters	127	127	128	128	128	128

NOTES: OLS coefficient estimates of Equation 1. The dependent variable is an index ranging between 0 and 3 that captures the intensity of sexual violence. Explanatory variables are perpetrator-specific ethnic characteristics related to descent, residence patterns, family arrangements, and subsistence activities. All explanatory variables are normalized and range between 0 and 1. Observations are at the perpetrator-victim-conflict level. Columns (1), (3) and (5) report coefficients for covariates weighted by the size of the ethnic group, while columns (2), (4) and (6) report coefficients for unweighted covariates. Standard errors are clustered at the perpetrator's level. *** (**) (*) indicate significance at the 1% (5%) (10%) level.

TABLE B-4: *Perpetrator's ethnic characteristics and sexual violence in armed conflict. Robustness tests*

	Dependent Variable: Sexual Violence (0-3)		
	Victim FE	Country FE	No time variation
	(1)	(2)	(3)
The Slave Trade			
Transatlantic Slave Trade	-6.01**	-1.09***	-1.09***
	(2.434)	(0.259)	(0.397)
Adjusted R^2	0.55	0.27	0.21
Indian Ocean Slave Trade	0.74**	0.81***	0.43
	(0.292)	(0.240)	(0.627)
Adjusted R^2	0.34	0.26	0.20
Mean dep. var.	0.62	0.62	0.55
Conflict fixed effect	✓		✓
Country fixed effect		✓	
Year fixed effect	✓	✓	
Conflict-specific time trend	✓		
Country-specific time trend		✓	
Victim fixed effect	✓		
Observations	880	900	196
Clusters	127	128	100

NOTES: OLS coefficient estimates of Equation 1. The dependent variable is an index ranging between 0 and 3 that captures the intensity of sexual violence. Explanatory variables are perpetrator-specific ethnic characteristics related to descent, residence patterns, family arrangements and subsistence activities. All explanatory variables are normalized and range between 0 and 1. Observations are at the perpetrator-victim-conflict level. Columns (1), (3) and (5) report coefficients for covariates weighted by the size of the ethnic group, while columns (2), (4) and (6) report coefficients for unweighted covariates. Standard errors are clustered at the perpetrator's level. *** (**) (*) indicate significance at the 1% (5%) (10%) level.

TABLE B-5: *Perpetrator's ethnic characteristics and sexual violence in armed conflict. Robustness tests*

	Dependent Variable: Sexual Violence (0-3)					
	Victim FE		Country FE		No time variation	
	Weighted	Unweighted	Weighted	Unweighted	Weighted	Unweighted
	(1)	(2)	(3)	(4)	(5)	(6)
Ethnic Gender Inequality Index						
eGII	0.84	0.81*	1.47***	1.48***	1.54**	1.59***
	(0.550)	(0.472)	(0.461)	(0.450)	(0.630)	(0.585)
Adjusted R^2	0.547	0.548	0.296	0.298	0.260	0.269
Restricted eGII	1.20	1.07*	1.01**	1.10***	1.18**	1.40***
	(0.726)	(0.594)	(0.458)	(0.391)	(0.566)	(0.528)
Adjusted R^2	0.550	0.550	0.283	0.287	0.251	0.267
Conflict fixed effect	✓	✓			✓	✓
Country fixed effect			✓	✓		
Year fixed effect	✓	✓	✓	✓		
Conflict-specific time trend	✓	✓				
Country-specific time trend			✓	✓		
Victim fixed effect	✓	✓				
Observations	880	880	900	900	266	266
Clusters	127	127	128	128	128	128

NOTES: OLS coefficient estimates of Equation 1. The dependent variable is an index ranging between 0 and 3 that captures the intensity of sexual violence. Explanatory variables are perpetrator-specific ethnic characteristics related to descent, residence patterns, family arrangements and subsistence activities. All explanatory variables are normalized and range between 0 and 1. Observations are at the perpetrator-victim-conflict level. Columns (1), (3) and (5) report coefficients for covariates weighted by the size of the ethnic group, while columns (2), (4) and (6) report coefficients for unweighted covariates. *** (**) (*) indicate significance at the 1% (5%) (10%) level.

TABLE B-6: *Restricted eGII: PCA loadings*

Variables	*Loading*
Gender Equal Traits	
Matrilineal	-0.29
Dependence on agriculture	-0.41
Gender Unequal Traits	
Virilocal	0.30
Dependence on pastoralism	0.57
Dependence on animal husbandry	0.57
Kaiser-Meyer-Olkin's measure of sampling adequacy	0.58

NOTES: Loadings from the principal component analysis on the restricted eGII.

TABLE B-7: *Cultural distance in gender norms and sexual violence in armed conflict. Robustness tests (I): controlling for victim's characteristics*

	Dep. Variable: Sexual Violence (0-3)			
	(1)	(2)	(3)	(4)
Victim's eGII	0.83	-1.14	1.13	
	(0.879)	(0.806)	(0.997)	
Perpetrator *more* unequal	2.77**		2.64**	2.17
	(1.057)		(1.047)	(2.101)
Perpetrator *less* unequal		-0.75	-0.49	0.88
		(0.658)	(0.641)	(1.527)
Conflict fixed effect	✓	✓	✓	✓
Year fixed effect	✓	✓	✓	✓
Conflict-Specific time trends	✓	✓	✓	✓
Victim fixed effect				✓
Mean Dep. Var	0.62	0.62	0.62	0.62
Observations	643	643	643	625
Adjusted R^2	0.60	0.59	0.60	0.70

NOTES: OLS coefficient estimates of Equations 3 and 4. The sample is restricted to include inter-ethnic conflicts only. The dependent variable is an index ranging between 0 and 3 that captures the intensity of sexual violence. The explanatory variables are the following: the victim's eGII (weighted by the ethnic group land area); the absolute distance in the eGII between perpetrator and victim when the perpetrator is more gender unequal than the victim; the absolute distance in the eGII between perpetrator and victim when the perpetrator is less gender unequal than the victim. Standard errors are clustered at the dyad level. *** (**) (*) indicate significance at the 1% (5%) (10%) level.

TABLE B-8: *Cultural distance in gender norms and sexual violence in armed conflict. Robustness tests (II): abstracting from temporal variation*

	Dependent Variable: Sexual Violence (0-3)			
	(1)	(2)	(3)	(4)
Absolute distance ($\lvert eGII_p - eGII_v \rvert$)	1.98*			
	(1.027)			
Perpetrator's eGII		-0.19	2.71	1.36
		(0.888)	(1.753)	(1.916)
Perpetrator *more* unequal		1.95*		2.26**
		(1.007)		(0.966)
Perpetrator *less* unequal			1.60	2.11
			(1.745)	(1.854)
Conflict fixed effect	✓	✓	✓	✓
Mean Dep. Var	0.54	0.54	0.54	0.54
Observations	189	189	189	189
Adjusted R^2	0.226	0.245	0.234	0.256

NOTES: OLS coefficient estimates of Equations 3 and 4. The sample is restricted to include inter-ethnic conflicts only. The dependent variable is an index ranging between 0 and 3 that captures the intensity of sexual violence. The explanatory variables are the following: the absolute distance in the eGII between perpetrator and victim; the perpetrator's eGII (weighted by the ethnic group land area); the absolute distance in the eGII between perpetrator and victim when the perpetrator is more gender unequal than the victim; the absolute distance in the eGII between perpetrator and victim when the perpetrator is less gender unequal than the victim. Robust standard errors are reported in parenthesis. *** (**) (*) indicate significance at the 1% (5%) (10%) level.

TABLE B-9: *Cultural distance in gender norms and sexual violence in armed conflict. Robustness tests (III): alternative fixed effects and alternative versions of the eGII*

	Dependent Variable: Sexual Violence (0-3)			
	Conflict-year FE	Country FE	Unweighted eGII	Restricted eGII
	(1)	(2)	(3)	(4)
Perpetrator *more* unequal	1.18**	1.62***	1.14*	1.27**
	(0.521)	(0.514)	(0.618)	(0.502)
Perpetrator *less* unequal	1.60	1.45	0.81	1.55
	(1.067)	(1.246)	(1.239)	(1.260)
Conflict fixed effect	✓	✓	✓	✓
Year fixed effect	✓	✓	✓	✓
Conflict-specific time trends		✓	✓	✓
Country fixed effect		✓		
Conflict-Year fixed effect	✓			
Perpetrator fixed effect	✓	✓	✓	✓
Mean Dep. Var	0.62	0.62	0.62	0.62
Observations	604	623	623	623
Adjusted R^2	0.64	0.74	0.74	0.74

NOTES: OLS coefficient estimates of Equations 3 and 4. The sample is restricted to include inter-ethnic conflicts only. The dependent variable is an index ranging between 0 and 3 that captures the intensity of sexual violence. The explanatory variables are the following: the absolute distance in the eGII between perpetrator and victim when the perpetrator is more gender unequal than the victim; the absolute distance in the eGII between perpetrator and victim when the perpetrator is less gender unequal than the victim. Standard errors are clustered at the dyad level. *** (**) (*) indicate significance at the 1% (5%) (10%) level.

TABLE B-10: *Cultural distance in gender norms and sexual violence in armed conflict. Robustness tests (V): assigning to governments a country-level measure of the eGII*

	Dependent Variable: Sexual Violence (0-3)				
	(1)	(2)	(3)	(4)	(5)
Absolute distance ($\lvert eGII_p - eGII_v \rvert$)	1.69***				
	(0.495)				
Perpetrator's eGII		0.59	1.70*	0.60	
		(0.832)	(0.886)	(0.934)	
Perpetrator *more* unequal		1.83**		1.83**	1.75***
		(0.867)		(0.867)	(0.434)
Perpetrator *less* unequal			-0.43	0.01	1.52
			(0.933)	(0.855)	(1.112)
Conflict fixed effect	✓	✓	✓	✓	✓
Year fixed effect	✓	✓	✓	✓	✓
Conflict-specific time trends	✓	✓	✓	✓	✓
Perpetrator fixed effect	✓				✓
Mean dep. var.	0.62	0.62	0.62	0.62	0.62
Observations	633	653	653	653	633
Adjusted R^2	0.600	0.367	0.360	0.366	0.599

NOTES: OLS coefficient estimates of Equations 3 and 4. The sample is restricted to include inter-ethnic conflicts only. Government forces' eGII is a country-level measure capturing the weighted average of ethnic groups' eGII within a country, weighted by the size of their land area. The dependent variable is an index ranging between 0 and 3 that captures the intensity of sexual violence. The explanatory variables are the following: the absolute distance in the eGII between perpetrator and victim; the absolute distance in the eGII between perpetrator and victim when the perpetrator is more gender unequal than the victim; the absolute distance in the eGII between perpetrator and victim when the perpetrator is less gender unequal than the victim. Standard errors are clustered at the dyad level. *** (**) (*) indicate significance at the 1% (5%) (10%) level.

TABLE B-11: *Cultural distance in gender norms and sexual violence in armed conflict. Robustness tests (IV): multi-way cluster*

	Dependent Variable: Sexual Violence (0-3)				
	(1)	(2)	(3)	(4)	(5)
Absolute distance ($\lvert eGII_p - eGII_v \rvert$)	1.53**				
	(0.623)				
Perpetrator's eGII		0.58	2.05	1.13	
		(0.468)	(1.398)	(1.379)	
Perpetrator *more* unequal		1.44**		1.51***	1.53**
		(0.584)		(0.569)	(0.650)
Perpetrator *less* unequal			0.20	0.64	1.56
			(1.328)	(1.350)	(1.488)
Conflict fixed effect	✓	✓	✓	✓	✓
Year fixed effect	✓	✓	✓	✓	✓
Conflict-specific time trends	✓	✓	✓	✓	✓
Perpetrator fixed effect	✓				✓
Mean dep. var.	0.62	0.62	0.62	0.62	0.62
Observations	623	643	643	643	623
Adjusted R^2	0.579	0.379	0.374	0.379	0.578

NOTES: OLS coefficient estimates of Equations 3 and 4. The sample is restricted to include inter-ethnic conflicts only. The dependent variable is an index ranging between 0 and 3 that captures the intensity of sexual violence. The explanatory variables are the following: the absoluthttps://www.overleaf.com/project/5db956fde60e2f00019d9daae distance in the eGII between perpetrator and victim when the perpetrator is more gender unequal than the victim; the absolute distance in the eGII between perpetrator and victim when the perpetrator is less gender unequal than the victim. Standard errors are clustered at the perpetrator and victim level. *** (**) (*) indicate significance at the 1% (5%) (10%) level.