GROUP INEQUALITY AND CIVIL CONFLICT

JOHN D. HUBER AND LAURA MAYORAL[1]

January 3, 2018

ABSTRACT

This paper explores empirically the relationship between group-level economic inequality and civil conflict. Previous theoretical research argues that high inequality within an ethnic group can make inter-ethnic conflict *more violent* because such inequality decreases the opportunity cost to poor group members of fighting and also decreases the opportunity cost to rich group members of funding the conflict. Drawing on a new data set using individual-level surveys to measure inequality at the ethnic group level, we provide strong evidence for this argument. Other studies have argued that economic inequality *between* ethnic groups – "horizontal inequality" – creates grievances that lead groups to *initiate* civil conflicts, and this research also provides empirical support for this argument. We find, however, that the empirical findings regarding horizontal inequality are not robust. Our study therefore underlines the value of focusing on the *capacity* of groups to fight rather than on their economic grievances if one wishes to limit the destruction of civil conflicts.

Keywords: Ethnicity, within-group inequality, horizontal inequality, civil conflict.

JEL: D63, D74, J15, O15

[1] This paper is a revised version of our previous work "Inequality, Ethnicity and Civil Conflict". Huber: Department of Political Science, Columbia University, jdh39@columbia.edu; Mayoral: Institut d'Anàlisi Económica, CSIC, Barcelona GSE, and University of Gothenburg; laura.mayoral@iae.csic.es. John Huber is grateful for financial support from the National Science Foundation (SES-0818381). Laura Mayoral gratefully acknowledges financial support from the Generalitat de Catalunya, and the Ministry of Economy and Competitiveness Grant number ECO2015-66883-P and from National Science Foundation Grant SES-1629370 (PI: Debraj Ray). We received helpful comments from Lars-Erik Cederman, Joan Esteban, Debraj Ray and seminar participants at various venues where this paper was presented. We also thank Sabine Flamand, Andrew Gianou and Tom Orgazalek for superb research assistance.

1. INTRODUCTION

The tragic civil war in Syria, which broke out in 2011 and which pits a range of ethnic and sectarian groups against the government, has cost hundreds of thousands of lives and displaced millions. While it is a particularly depressing case, Syria joins a growing list of countries that have suffered from serious intrastate conflicts, which have replaced inter-state wars as the nexus for large scale violence in the world.[2] Conflicts, however, vary a great deal in their intensity, with some conflicts leading to large numbers of fatalities and others to a small number. Over the period 1946-2009, for example, the total number of battle-related deaths in the 25% of least-deadly civil conflicts is less than 5000 people in total (Lacina and Gleditsch, 2005). By contrast, the equivalent figure in the top 25% of most-deadly conflicts is close to 5 million people.[3] Thus, if one wishes to limit the death and destruction of civil conflict, one must understand not simply the triggers of conflict. It is perhaps even more important to focus on the factors that influence a conflict's severity.

This paper examines empirically the role of economic inequality in both the onset and severity of civil war. The analysis has two features worth noting. First, it is at the ethnic group level. Much work on economic inequality and civil conflict has been influenced by long-standing arguments by Karl Marx, Dahrendorf (1959) and Gurr (1970, 1980), who argue that inequality breeds grievances which in turn lead to civil violence. However, national level studies have not found empirical support for this idea, leading some to conclude that economic grievances do not cause conflict (e.g., Lichbach 1989, Fearon and Laitin 2003 and Collier and Hoeffler 2004). But most civil conflict is rooted in ethnic, sectarian or other identity groups (Doyle and Sambanis 2006, Fearon and Laitin 2003), making it important to consider arguments about how economic inequality measured with respect to groups affects the propensity of groups to participate in civil wars (e.g., Cederman, Weidmann and Gleditsch 2011, Cederman, Gleditsch and Buhaug 2013).[4]

[2]Gleditsch et al. (2002), for example, find that since WWII, there were 46 interstate conflicts with more than 25 battle-related deaths per year, 22 of which have killed at least 1,000 over the entire history of the conflict. Over the same period, there were 181 civil conflicts with more than 25 battle-related deaths per year, and almost half of them had killed more than 1,000 people.

[3]These figures are based on the "best estimate" of the Battle Deaths dataset, version 3.1; see Lacina and Gleditsch (2005).

[4]See Esteban, Mayoral and Ray (2012) and Arbatli, Ashraf and Galor (2015) for recent evidence on the connection between ethnic structure and conflict.

Second, the analysis distinguishes between arguments about group-based economic inequality and the *onset* of civil conflict, on one hand, and group-based inequality and the *intensity* of civil conflict, on the other. With respect to conflict onset, a prominent line of research argues that horizontal inequality – economic differences *across* groups – leads groups to start civil wars (e.g., Stewart 2002, Cederman et al. 2011). Such inequality exists when a group is particularly poor or particularly rich, so relatively poor and rich groups should be most likely to initiate conflicts. These arguments about horizontal inequality and civil war onset, however, have little to say about the intensity of fighting. Small and powerless groups might spark a conflict but then be quickly crushed by the government, or they might keep low intensity conflicts simmering over time. Arguments about inequality and the intensity of civil wars have therefore focused on a different dimension of inequality. In particular, recent theoretical research argues that internally unequal groups are most likely to engage in severe conflicts because such groups have both the labor (poor people) and the capital (rich people) necessary to fight (Esteban and Ray, 2011).

This paper makes three contributions to studies of inequality and civil war. First, we introduce a new data set that uses surveys to measure group inequality and argue that these measures present important advantages over estimates based on geo-coded data, particularly when measuring within-group inequality. Second, we use this data to provide evidence of a robust association between within-group inequality and the *intensity* as well as the *incidence* of civil war. To the best of our knowledge, this is the first paper that analyses empirically the link between within-group inequality and conflict *intensity* at the group level.[5] The estimated association – based on cross-national regressions which limit the possibility of making causal claims[6] – is large and robust. Also, consistent with theoretical arguments emphasizing the role of income heterogeneity in the strength of rebel fighters, our results also show that the association between within group-inequality and conflict *onset* is weaker.

[5]Previous studies have focused instead on conflict onset, which as noted above, is not clearly linked to the ER theory. Østby, Nordås and Rød (2009) find a positive association between *within-region* inequality and conflict onset in 22 countries in Sub-Saharan Africa, and Kuhn and Weidmann (2015), which we discuss below, find a positive association between within-group inequality and conflict onset, using a global dataset that has the ethnic group – rather than the region– as unit of analysis.

[6]See Section 5, however, which is devoted to assessing issues of reverse causation and omitted variable bias.

Third, using both the survey-based measures and geo-coded data from previous studies, we provide evidence that previous empirical claims that horizontal economic inequality is related to conflict onset are not robust. Indeed, we find essentially no evidence that the two variables are related. We argue that this null finding may be unsurprising given that from a theoretical point of view it is not clear why the relative wealth of a group should be systematically related to its propensity to start a war, as has also been pointed out in the literature (Esteban and Ray, 2011).

The picture our paper paints regarding the role of inequality in civil conflict is thus quite different than previous research. In contrast to group level research on conflict *onset* but consistent with national-level research on overall economic inequality, the analysis provides no support for the argument that economic grievances spark civil wars. But in contrast to the national level research dismissing the role of economic inequality, we find that inequality can play an important part in understanding civil conflict, albeit one unrelated to grievances or conflict onset. Instead, unequal groups are associated with a higher capacity to fight and therefore with more severe civil conflicts. Our findings therefore suggest that efforts to reduce the severity of civil conflicts should focus on situations where groups have the highest capacity to fight, and that inequality within a group is central in this regard.

The paper is organized as follows. Section 2 reviews the existing arguments that represent the point of departure for our analysis, discussing first the relationship between intra-group inequality and conflict intensity and then discussing horizontal inequality and conflict onset. Section 3 introduces a new data set that uses individual-level surveys to measure between- and within-group inequality. The empirical analysis is found in the three sections that follow. Section 4 focuses on conflict intensity and incidence as dependent variables, followed in Section 5 by a discussion of the possibility that results are due to omitted variable bias or reverse causality. Section 6 then turns to empirical analysis of civil war onset. Section 7 concludes. Appendix A provides additional analyses of the data, and (on-line) Appendices B and C contain materials regarding the data and its construction.

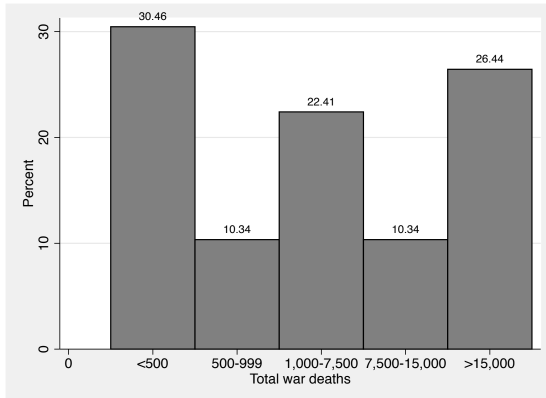


Figure 1. Variation in the qualitative nature of civil wars

2. Ethnic inequality, conflict onset and conflict intensity

Although empirical research on civil conflict typically focuses on why wars break out, it is important to recognize that the damage wrought by civil wars varies a great deal (Lacina 2006). Figure 1 summarizes some of these differences for 174 civil conflicts with battle death information in the PRIO data (see Lacina and Gleditsch 2005). Around one third of wars have less than 500 total casualties, and 40% have less than 1,000. But some wars lead to a great deal of death, with 26% of wars having more than 15,000 casualties. Thus, limiting the devastation from civil conflicts requires that we understand the factors that can lead conflicts to become particularly severe.

To this end, it is important to recognize that group resources can play different roles in the initiation and intensity of conflicts. If the amount of resources available to a group is large, conflicts involving this group can be expected to severe, creating a deterrent to conflict onset.[7] But if a group does become involved in a conflict, when the group has substantial resources the intensity of conflict should be high. Thus the type of economic inequality that affects conflict onset might differ from the type that facilitates conflict intensity. More generally, the onset and the intensity of civil war should be explained by different theoretical frameworks and should rest on different empirical foundations, as recently emphasized Bazzi and Blatmann (2014). This section describes how two different dimensions of ethnic income heterogeneity – *between* and *within*-group inequality– have

[7] In a related vein, Cunningham, Gleditsch and Salehyan (2009) find that governments settle quickly and make greater concessions when rebels are strong. See also Fearon (2004).

been linked in previous theoretical arguments to the two dimensions of ethnic civil conflict. We begin with within-group inequality, which has received less attention in the literature.

2.1. **Within-group inequality and conflict intensity.** Esteban and Ray (2011) (henceforth "ER") provide a simply theory linking one aspect of inequality – that which exists *within* groups – to conflict intensity. They argue that for a group to fight effectively, it needs financial resources and labor (i.e, fighters). Sustaining a high-intensity conflict therefore has at least two opportunity costs: the cost of contributing resources and the cost of contributing one's labor to fight. Inequality within a group simultaneously decreases both opportunity costs. When the poor within a group are particularly poor, they will require a relatively small compensation for fighting, and when the rich within a group are particularly rich, the opportunity cost of resources to fund fighters will be relatively low. Thus, groups with high income inequality should have the greatest capacity to wage destructive civil conflicts. While ER's theory emphasizes the role of within-group inequality("WGI") in fueling conflict, their model assumes that a conflict already exists and thus is silent about the relation between conflict onset and income heterogeneity within the group. This relationship is likely to be more ambiguous: high within-group inequality may facilitate the mobilization of combatants, but the threat of a highly destructive conflict could also deter conflict onset in the first place by encouraging negotiation and compromise by the government.

ER (2011) provide examples from Africa, Asia and Europe to illustrate the causal mechanisms in their theory. In their survey of the literature on ethnic conflict, Fearon and Laitin (2000) also describe examples where the elites promote ethnic conflict and combatants are recruited from the lower class to carry out the killings, including Bosnia (the "weekend warriors," a lost generation who sustained the violence by fighting during the weekends and going back to their poorly-paid jobs in Serbia during the week), Sri Lanka (where gang members fought on the Sinhalese side), and Burundi. They conclude,

> [O]ne might conjecture that a necessary condition for sustained *ethnic violence* is the availability of thugs (in most cases young men who are ill-educated, unemployed or underemployed, and from small towns) who can be mobilized by nationalist ideologues, who themselves, university educated, would shy away from killing their neighbors with machetes. (p. 869)

The case of the Rwandan genocide is also suggestive. In the spring of 1994, the Hutu majority carried out a massacre against the Tutsi minority where 500,000 to 800,000 Tutsi and moderate Hutus that opposed the killing campaign were assassinated. In the years immediately prior to the genocide, Rwanda suffered a severe economic crisis motivated by draughts, the collapse of coffee prices, and a civil war. Verwimp (2005) documents an increase in within-group inequality among the Hutu population prior to the genocide: on the one hand, a sizable number of households that had been middle-sized farmers lost their land and became wage workers in agriculture or low skilled jobs. On the other, rich farmers with access to off-farm labor were able to keep and expand their land. This new configuration encouraged the Northern Hutu elites to use their power to instigate violence. Backed by the Hutu government, these elites used the radio (particularly RTLM) and other media to begin a propaganda campaign aimed at fomenting hatred of the Tutsis by Hutus (Yanagizawa-Drott, 2012). The campaign had a disproportionate effect on the behavior of the unemployed and on delinquent gang thugs in militias throughout the country, who had the most to gain from engaging in conflict and the least to lose from not doing so (Melvern 2000). Importantly, the campaign made it clear that individuals who engaged in the ethnic-cleansing campaign would have access to the property of the murdered Tutsi (Verwimp, 2005). Thus, the rich elites "bought" the services of the recently impoverished population by paying them with the spoils of victory, something that was more difficult to undertake prior to the economic crisis.[8]

[8] It is also worth noting that micro-level studies of participation emphasize that richer elites recruit the poor to fight. Brubaker and Laitin (1998), for example, argue that most ethnic leaders are well-educated and from middle-class backgrounds while the lower-ranking troops are more often poorly educated and from working-class backgrounds. In their study of Sierra Leone's civil war, Humphreys and Weinstein (2008) find that factors such as poverty, a lack of access to education, and political alienation are good predictors of conflict participation and that they may proxy, among other factors, for a greater vulnerability to political manipulation by elites. Justino (2009) also emphasizes that poverty is a leading factor in explaining participation in ethnic conflict.

Although the ER theory provides a clear rationale for why intra-group inequality should be positively related to the incidence of conflict, the opposite argument has also been made. Such intra-group inequality could create resentment among the poor and reduce group cohesiveness (Sambanis and Milanovic, 2011) which in turn could have a negative impact on conflict.Thus, it is important to understand empirically the relationship between within-group inequality and conflict intensity.

2.2. **Horizontal inequality and conflict onset.** Substantial previous research links between-group inequality – often referred to as horizontal inequality – to the onset of conflict. When a group has large economic differences with other groups, the argument goes, it should be more likely to *initiate* conflict regardless of whether the group is relative rich or relatively poor. For poorer ethnic groups, the concept of relative deprivation is central, with such groups having incentives to initiate conflict in order to gain resources, reducing the feelings of alienation and grievance (e.g., Cederman, Weidmann and Gleditsch 2011, Cederman, Gleditsch and Buhaug 2013, Stewart 2000 and Stewart 2002).[9] The larger the income gap between the groups, the greater the potential prize, and hence the greater the incentive for conflict by the poorer group (e.g., Acemoglu and Robinson 2005, Cramer 2003, Stewart 2002 and Wintrobe 1995). The fact that inequalities are linked to ethnic identities is important because identity-based inequalities are easily politicized by elites, with such politicization typically involving a framing strategy where an afflicted group blames one or more other groups for injustice. Additionally, a central problem in initiating conflict is mobilization, and ethnic identities are held to facilitate solutions to the collective action problem associated with waging civil war. Thus, relatively poor ethnic groups are more likely to initiate civil violence.

This is only one-half of the story about horizontal inequality and ethnic civil conflict. Rich groups should be concerned that the government will expropriate their wealth to redistribute it in the rest of the country, giving such groups incentives to engage in pre-emptive attacks and/or secession wars to diminish threats against them. Relatively rich and poor groups should therefore

[9] For a detailed theoretical account of the link between horizontal inequality and civil conflict, as well as a review of the relevant literature, see Cederman, Gleditsch and Buhaug (2013, chapter 3).

be most likely to initiate conflict, creating a link between horizontal inequality and civil war onset (e.g., Cederman, Gleditsch and Buhaug 2013, pp. 97-98, and Cederman, Weidmann and Gleditsch 2011, p. 478). A number of empirical studies, which we discuss in Section 6 below, provide support for the horizontal inequality argument.

Several observations about the horizontal inequality argument deserve highlighting. First, if one considers the strategic incentives of groups, it is not clear that a group should be most likely to initiate civil war when it is relatively rich or poor. A relatively poor group is less likely to succeed in conflict because it lacks the resources necessary for success. If a very poor group is going to be crushed when it initiates conflict, it is reasonable to ask why it will attack in the first place. A rise in the income of a group might even enhance its capacity to fund militants and thus its probability of success in conflict. As a result, the closing of the income gap between two groups – rather than its widening – might ignite conflict.[10] This argument is related to the so-called Thucydides's Trap (Allison 2017), which states that when a disadvantaged group becomes more powerful and threatens to displace a ruling one, war becomes more likely. Mitra and Ray (2014) present supportive evidence from the Muslim-Hindu conflict in India, showing that when the gap between the income of those groups narrows, conflict intensifies. This point is also compatible with the evidence showing that economic modernization might facilitate rather than discourage ethnic conflict (Tellis, Szayna, and Winnefeld 1998, Chua 2003). Finally, using a logic similar to ER (2011), Adhvaryu et al. (2017) use data from Africa to show that the probability of conflict is low when the parties involved are poor (so horizontal inequality is low), it increases when one of the parties is relatively rich and the other is poor, and it is highest when both groups are rich (so horizontal is again low).

Similar strategic considerations should apply to rich groups: if rich groups understand that they will be difficult to defeat, it is not clear why relatively rich groups will need to undertake preemptive attacks. Indeed, if we introduced uncertainty about the relative military strength of rich and poor groups, we might expect that conflict will be highest when economic differences are most modest, as these will be the cases where the relatively poor group has the strongest belief it

[10]See Esteban and Ray (2011) for a discussion of this issue.

can win and the relatively rich group has the most to gain from pre-emption. While it is not our purpose to develop a theory about how economic grievances interact with strategic considerations regarding expectations of winning, we would underline that these strategic considerations work against the arguments about horizontal inequality and are notably absent from them.

Finally, the theoretical logic linking poor groups to the outbreak of civil conflict could be quite different from that posited in horizontal inequality arguments about grievance. As described above, HI arguments emphasize that poor groups have the most to *gain* from winning a conflict. But research on poverty and civil conflict emphasize instead that poorer groups have *the least to lose* (e.g., Collier and Hoeffler 2004). Since the opportunity cost of fighting for poor people is relatively low, poorer groups might be more likely to initiate conflict because their labor can be bought rather cheaply.[11] Although both the inequality and opportunity cost perspectives suggest poorer groups should be more likely to initiate conflict, the difference between the two types of arguments comes from thinking about rich groups. From the "poverty" perspective, relatively rich groups will not engage in conflict; in fact, they would have the highest opportunity cost of doing so. From the "horizontal inequality perspective," both relatively rich and relatively poor groups have incentives to initiate conflict.

Table 1. The Relation between Ethnic Inequality and Conflict

	Conflict Onset	Conflict Intensity
Horizontal inequality	**Positive** Poor (Rich) groups start conflicts to gain (preserve) resources (Cederman et al., 2011)	Ambiguous
Within-group inequality	Ambiguous	**Positive** Unequal groups have labor and capital necessary to sustain intense conflicts (Esteban and Ray, 2011)

Note. This table summarizes the two most prominent strands of research on group-based inequality and its connections with conflict initiation and intensity.

The two strands of research on group-based inequality are summarized in Table 1. Arguments about inequality between groups focus on what groups have to gain or lose, and these perceived

[11] See also Dal Bó and Dal Bó (2011), Dube and Vargas (2014) and Bazzi and Blattman (2015).

gains and losses are independent of whether a group can actually sustain a fight. This research therefore focuses on the link between horizontal inequality and the outbreak of conflict but it do not lead to clear expectations regarding the intensity of conflict. By contrast, arguments about inequality within groups does not provide clear expectations about conflict onset. High within-group inequality, for example, may facilitate the mobilization of combatants, but this could in fact deter conflict onset by facilitating negotiation and compromise by the government. But the theoretical arguments about within-group inequality do lead to a clear expectation about which types of groups can wage conflicts most effectively once conflicts begin.

3. Measuring ethnic inequality at the group-level

Previous research on horizontal inequality has typically relied on geo-coded data of economic activity, along with assumptions about the spatial location of groups. As we discuss below, however, it is very problematic to use spatial data to measure inequality *within* groups. This section therefore begins by describing how we use individual-level surveys to measure group-level inequality.[12] We then discuss the data used in previous research. Finally, we discuss the main strengths and weaknesses of using geo-coded versus surveys data to measure group inequality.

3.1. **Measuring group-based inequality using surveys.** In order to have a broad cross-national sample of countries, we include a variety of surveys to measure individual income and ethnic identity. These surveys have quite heterogenous measures of respondent income. The very best possible available surveys, which we call "HES" (for "Household Expenditure Survey"), have the most careful measures of household income and expenditures in existence. These include the Luxembourg Income Study, the Living Standards Monitoring Surveys, other similar household expenditure surveys, and a handful of national censuses. Ideally one would use only such surveys, but the problem is that there are simply not enough of them, particularly when one considers the need to measure ethnic identity. We therefore also use a variety of other types of surveys that have less precise measures of income. These include the World Values Surveys (WVS), which typically

[12]Previous research has used surveys to aggregate information across groups into country-level measures. Østby (2008), for example, develops various country-level measures of horizontal inequality and polarization that are useful in studies where a geographic space like a country or region (rather than the group) is the unit of analysis.

has about 10 household income categories per country, and the Comparative Study of Elections Surveys (CSES), which reports income in quintiles. We also include DHS and Afrobarometer surveys that do not have household income data, but rather have information on various assets that households possess. These asset variables are used to measure "income," as described in Appendix B (available on-line), which contains details about the data and the data construction process.

To measure the ethnic identify of respondents, we rely on the definitions of ethnic groups from Fearon (2003) and from Cederman, Wimmer and Min (2010). Appendix B provides further details regarding the mapping from survey categories to groups. The resulting data set, based on 232 surveys from the 1992-2008 time period, covers 446 groups from 89 countries. Appendix B provides further details about data coverage and includes a map of the countries with useful surveys, as well as a list of all surveys.

The survey information on ethnicity and "income" can be used to compute group-level income and Gini coefficients, but since the surveys measure respondent "income" in different ways, we face the problem of survey comparability. This is a standard challenge: the observations in Deininger and Squire's classic (1996) data set, for instance, differ in many respects (most significantly, in their income definitions and their reference units) and are rarely comparable across countries or even over time within a single country. Its successor, the World Income Inequality Database (WIID), faces identical challenges. Thus, if scholars wish to conduct cross-national research on inequality using such measures, they must adopt methodologies to adjust the measures from different surveys to make them comparable.

We consider two approaches. The *Ratio approach*, which yields a variable we call G^R, is similar to the methodology employed by Solt (2009) to construct the Standardized World Income Inequality Dataset (SWIID).[13] This approach uses external data on the Gini – the SWIID – to construct country-level adjustment factors based on the ratio of country-level Ginis from surveys to SWIID Ginis. These adjustment factors based on the ratios are then applied to the group Ginis to obtain the adjusted group Ginis. The second procedure is the *Intercept approach*, which shares the

[13]The SWIID provides comparable (country-level) Gini indices of gross and net income inequality for 173 countries from 1960 to the present and is one of the most thorough attempts to tackle the comparability challenge.

same spirit as the original Deininger and Squire (1996) exercise. To remove average differences due to varying survey methodologies, this approach first calculates the Gini coefficient for each group using the surveys and then regresses these Ginis on survey, time and country dummies, with HES as the omitted category (because the HES surveys are the best-available estimates of income distributions in the world). The coefficients on the survey dummies are then used to adjust the group inequality measures to remove average differences that could be traced to different survey types. The intercept approach yields a variable we call G^I. Since the time period is relatively short, we often (but not always) define our measures of group Gini as the average of the adjusted group Ginis from all the available surveys in a country. Appendix B shows that there is a strong correlation between G^I and G^R. It also describes how the surveys are used to measure group GDP, which is an input for the survey-based measures of horizontal inequality.

3.2. **Measuring horizontal inequality.** To measure horizontal inequality, we calculate the difference between the economic well-being of a group and the economic well-being of other groups. Let G_g be the GDP per capita of group g and $\bar{G}_g$ be GDP per capita of the population not in group g.[14] The measure of the difference between a group's economic well-being and the economic well-being of other groups is simply

$$\text{HI(ABS)} = |G_g - \bar{G}_g|.$$

To attenuate the skewness of this variable and the influence of outliers, our main focus will be on HI(LN), which is the natural log of HI(ABS). The variable can be measured using the geo-coded data or using the survey data.

These two measures differ from those used in previous research on horizontal inequality. Appendix B.3 describes these previous measures and discusses their limitations in capturing the underlying concept of horizontal inequality. In what follows, although our primary measure of

[14]More specifically, $\bar{G}_g$ is defined as total GDP of all groups minus total GDP of G divided by total population in the country minus the population of g. If GDP per capita at the group level is available, $\bar{G}_g$ can also be computed as follows. If there are n groups with sizes $\pi_1 \ldots \pi_n$ and per capita GDPs, $G_i \ldots G_n$, them $\bar{G}_g = \frac{\sum_{i \neq g} \pi_i G_i}{(1-\pi_g)}$.

horizontal inequality will be HI(LN), we will also estimate models using measures from previous research.

3.3. **Survey vs. spatial data in measuring WGI and HI.** The dominant approach in previous group-level studies relies not on survey data, as we do here, but rather on geo-referenced data on the geographic location of ethnic groups, along with geo-referenced estimates of economic development.[15] To elaborate measures of horizontal inequality, the spatial data on group locations is linked to spatial data on economic output, for example using Nordhauss (2006) G-Econ data set (the approach taken by Cederman, Weidmann and Gledistch 2011, CWG henceforth), or using satellite images of light density at night to produce measures of group GDP per capita.

Using either survey or spatial data to measure group inequality can invite biases in measurement. Appendix B.5 addresses this issue. With respect to the validity of survey-based measures, the appendix makes three points: (1) although the survey data exist for only 89 countries, much fewer than with geo-coded data, these countries are quite representative of a much broader set of countries; (2) the surveys do a very good job of capturing the distribution of individuals across groups;[16] and (3) the measurement of individual income using surveys leads to measures of group GDPs that are closely related to those obtained from the geo-coded approach.

The appendix also describes several avenues by which the use of geo-coded data can lead to bias: (a) geocoded ethnic homeland boundaries are often inaccurate; (b) the approach requires difficult assumptions in cases where individuals from different groups reside in the same location;[17] (c) urban areas must be excluded, dismissing an important source of within-group inequality, and (d) while the spatial estimates of group inequality are linked to the ethnic home boundaries, EPR

[15]There are a number of data sets on the geographic location of groups, including the GREG, the GeoEPR and the Ethnologue. The GREG dataset (Weidmann, Rob and Cedarman 2010) is based on the Soviet Atlas Narodov Mira. The *Ethnologue* provides information on the spatial location of linguistic groups in much of the world. The group-level studies of horizontal inequality and conflict have relied on the GeoEPR dataset, described in Wucherpfennig et al. (2011), which utilizes an expert survey to determine the identity and location of politically relevant ethnic groups (Wimmer et al., 2009).

[16]The ELF index constructed from the surveys is extremely highly correlated with the ELF index constructed using the Fearon data.

[17]Regions with mixed ethnic composition are quite common, and Morelli and Rohner (2015) link this segregation itself to civil conflict.

conflict definition is agnostic as to where conflict takes place in the country, which can lead to measurement error in countries where groups are not highly regionally segregated.

The disadvantages of using geo-coded data are particularly severe when used to compute within-group inequality, which requires information about the income distribution of each ethnic group (not simply about the group average income, as in the case of the HI measures). As a result, there is little previous research that measures inequality within groups for a large number of countries.[18] A notable exception is Kuhn and Weidmann ("KW", 2015), who construct a global dataset of group Gini coefficients using nightlight emissions. To do this, they first divide ethnic homelands into cells of equal size (about 10×10 km), focusing on non-urban areas.[19] They then compute nightlight emissions per capita for each cell and all cells occupied by a group are used as inputs (i.e., as if they were data on individuals) to calculate the group's Gini coefficient.

A central disadvantage of this approach is that the measure is very sensitive to the cell size selected, which is arbitrary. Another important concern is the fact that urban areas are not considered in the computation of the measures. With over half the world's population living in urban areas (Angel 2012), the fact that urban cells are discarded introduces a potential source of bias, as an important source of within-group inequality (rural-urban inequality) is dismissed. Finally, the approach results in the loss of data for a considerable number of groups, and we do not know how characteristics of the groups for which measures do not exist are related to inequality or conflict. Further details are in Appendix B.5.

4. Empirical analysis of group inequality and the intensity of conflict

We now turn to the substantive focus of our paper, which is to analyze empirically the relationship between the group-level inequality measures and the measures of group involvement in civil conflict, beginning in this section with a focus on conflict intensity/incidence and then moving on to

[18] Østby, Nordås and Rød (2009), for example, use survey data from the Demographic Health Surveys in 22 countries in Sub-Saharan Africa. Their study calculates the Gini coefficient for each region and their analysis finds that regions with higher levels of inequality are most likely to experience the onset of conflict. Fjelde and Østby (2014) provide a similar regional-level study, focusing on civil unrest rather than civil war.

[19] KW point out that groups might be relatively geographically segregated in the countryside, this is unlikely to be the case in urban areas. Thus including urban cells can introduce measurement error.

the onset of civil conflict in Section 6.[20] To measure group involvement in conflict in a given year, we use the Ethnic Power Relations data set (EPR, Cederman et al. 2009).[21] Ethnic groups are coded as engaged in conflict if a rebel organization involved in the conflict expresses its political aims in the name of the group, and a significant number of members of the group participate in the conflict (see Wucherpfennig et al. 2012 for details). Following previous literature, we drop from the sample monopoly and dominant groups, as these groups represent the state and therefore cannot fight against themselves.[22]

The theory in Esteban and Ray (2011) argues that within-group inequality affects a group's capacity to wage and sustain conflict. Our primary dependent variable for testing this argument is INTENSITY, which captures conflict incidence and severity. The variable takes the value 0 if a group is at peace in a given year, the value 1 for each year in which an ethnic group is involved in armed conflict against the state resulting in more than 25 battle-related deaths, and the value 2 if the resulting number of battle deaths that year is larger than 1000. We also use four alternative dependent variables: INCIDENCE, which is a binary measure that takes the value 1 for each year in which an ethnic group is involved in armed conflict against the state resulting in more than 25 battle-related deaths, CONFLICT-SHARE is the share of years a group has been in conflict in the period 1992- 2010, as well as BATTLE DEATHS (BEST) and BATTLE DEATHS (LOW), which are the log of the number of battle related deaths using the best and low estimate from Lacina and Gleditsch (2005) (see Appendix A.1).

[20]To the best of our knowledge, this is the first paper that analyses empirically the link between within-group inequality and conflict *incidence* and *intensity* at the group level. Previous studies have focused instead on conflict onset, which as noted above, is not clearly linked to the ER theory. Østby, Nordås and Rød (2009) find a positive association between *within-region* inequality and conflict onset in 22 countries in Sub-Saharan Africa, and Kuhn and Weidmann (2015), which we discuss below, find a positive associatin between within-group inequality and conflict onset, using a global dataset that has the ethnic group – rather than the region– as unit of analysis.

[21]The data are accessed through the ETH Zurich's GROWup data portal (`http://growup.ethz.ch`). We combine data from the Ethnic Power Relations (EPR) Core Dataset 2014 (https://icr.ethz.ch/data/epr/core/), the ACD2EPR 2014 dataset on conflict (https://icr.ethz.ch/data/epr/acd2epr/) and the GeoEPR 2014 data set on group attributes, including group economic well being (https://icr.ethz.ch/data/epr/geoepr/). See Appendix B for further details.

[22]A group is classified as monopoly if the elite members hold monopoly power in the executive to the exclusion of members of all other ethnic groups. A group is classified as dominant if elite members of the group hold dominant power in the executive but there is some limited inclusion of "token" members of other groups who however do not have real influence on decision making.

Our baseline specifications estimate ordered logits given by

$$P(\text{INTENSITY}_{c,g,t} = i \mid \mathrm{X}_{c,g,t}) = P(\alpha_i \leq y^*_{c,g,t} \leq \alpha_{i+1} \mid \mathrm{X}_{c,g,t}) \quad i = \{0, 1, 2\},$$

where $\mathrm{X}_{c,g,t}$ is a vector of right-hand side variables and $y^*_{c,g,t}$ is a latent variable measuring conflict intensity defined as

$$\begin{aligned} y^*_{c,g,t} &= \mathrm{X}'_{c,g,t}\Lambda + \epsilon_{c,g,t} \\ &= \beta_1 \mathrm{G}^R_{c,g} + \beta_2 \mathrm{HI(LN)}_{c,g,t} + Z'_{c,g,t}\delta + W'_{c,t}\gamma + V'_c\eta + T'_t\gamma + y^*_{c,g,t-1} + \epsilon_{c,g,t}, \end{aligned}$$

where $Z_{c,g,t}$, $W_{c,t}$, V_c and T_t are vectors containing group-level controls, country-level controls and country and year dummies, respectively. Thus, all empirical models include (unreported) country and year indicator variables as well as a lagged dependent variable (to capture conflict persistence). Our regressions cover the period 1992-2010 and are based on 86 countries and 335 groups (at most). As explained in Section 3.1 and Appendix B.1, our inequality variables are based on surveys from the 1992-2008 time period. In our baseline regressions, WGI variables are time invariant (and obtained by averaging all the observations available for a group). We also compute alternative measures of WGI and show that our results are robust (see Appendix A). Control variables are described as they appear in the text. Detailed definitions of all variables as well as a table of summary statistics are provided in Appendix C.

Table 2 presents 7 models where we regress INTENSITY on G^R using ordered logit models with standard errors clustered at the country level. Model 1 includes only G^R, with no other controls (except the country and year fixed effects and the lagged dependent variable). The coefficient for G^R is positive and significant at the 1% level.[23] Model 2 includes HI(LN), our preferred measure of horizontal inequality, computed using survey data. Its coefficient is positive but insignificant while that of G^R remains largely unchanged. Column 3 adds four group-level controls: GROUP

[23] The adjusted inequality measures are generated regressors and therefore standard errors that do not take this fact into account are generally invalid since they ignore the sampling variation in such regressors. Nevertheless, for the purpose of testing whether the inequality variables are significantly different from zero, the sampling variation in the generated regressors can be ignored (at least asymptotically). See Wooldridge (2002), chapters 6 and 12 for additional details.

ELEV(SD) is the standard deviation of the elevation of the ethnic homeland; GROUP SIZE is the group's relative size; GROUP DIAMONDS and GROUP OIL are dummies indicating whether the ethnic homeland contains diamonds and oil, respectively (obtained from Prio-Grid, see Tollefsen et al. 2012). Column 4 introduces GROUP GDP, the (lagged) log of the per capita GDP of the group (computed using survey data as described in Section C.17), while Column 5 includes two time-varying country level controls, POP is the (lagged) log of the country's population and GDP is the log of the country's GDP per capita (lagged), both from the Penn World Tables (PWT). Column 6 adds two political controls: XPOLITY is a country's (lagged) democracy score based on Polity IV and EXCLUDED is a dummy capturing whether the group is excluded from power.[24]

Our conclusions regarding the inequality variables remain very stable when these controls are introduced: G^R is always significant and the 5% level and HI(LN) is never significant. Among the control variables, only XPOLITY is significantly associated with INTENSITY. Finally, column 7 drops GROUP GDP from model 6, since there is a high degree of correlation between the group and country-level GDP variables (.96). Column 7 shows that when group GDP is excluded, the coefficient of G^R *decreases* in magnitude but the level of significance remains identical.[25]

The coefficients for G^R are not only precisely estimated, they are substantively large. Using the estimates from column 7 in Table 2 (which has the smallest coefficient for G^R), we find that moving from the median value of G^R (.38) to the value in the 90th percentile (.54), while holding all other variables at their means, increases the predicted probability of conflict (i.e, the probability of observing strictly positive values of INTENSITY) by 262%. To put this effect into perspective, we have performed similar calculations for the other variable that is significant in Table 2, XPOLITY. In this case, moving from the median value of XPOLITY (4) to the value in the 90th percentile (7) while holding all other variables at their means reduces the predicted probability of conflict by 32%, quite small in comparison with the effect of G^R.

[24]XPOLITY combines 3 out of the 5 components of Polity IV and leaves out the two components (PARCOMP and PARREG) that are constructed using political violence in their definition (Vreeland 2008).

[25]Dropping GDP from column 6 and keeping GROUP GDP produces essentially identical results, see Appendix A.

Table 2. Within group inequality and the intensity of conflict

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
G^R	5.885**	5.575**	5.812**	6.230**	7.234**	7.756**	5.592**
	(0.019)	(0.027)	(0.014)	(0.042)	(0.028)	(0.030)	(0.016)
HI(LN)		0.103	0.002	-0.005	0.025	-0.010	0.013
		(0.531)	(0.991)	(0.978)	(0.893)	(0.961)	(0.955)
GROUP SIZE			-1.715	-1.842	-2.048	-1.140	-0.658
			(0.415)	(0.367)	(0.366)	(0.656)	(0.778)
GROUP ELEV. (SD)			-0.001	-0.001	-0.001	-0.001	-0.000
			(0.553)	(0.494)	(0.411)	(0.549)	(0.771)
GROUP DIAMONDS			1.055	1.046	1.166	0.925	0.927
			(0.344)	(0.342)	(0.290)	(0.407)	(0.416)
GROUP OIL			0.429	0.425	0.474	0.310	0.303
			(0.463)	(0.465)	(0.434)	(0.634)	(0.639)
GROUP GDP				0.177	0.660	0.900	
				(0.782)	(0.348)	(0.250)	
POP					2.874	-1.653	-1.603
					(0.456)	(0.477)	(0.491)
GDP					-1.215	0.098	0.938
					(0.328)	(0.955)	(0.520)
XPOLITY						-0.123*	-0.123*
						(0.062)	(0.062)
EXCLUDED GROUP						0.892	0.832
						(0.196)	(0.241)
INTENSITY(LAG)	3.938***	3.932***	4.120***	4.124***	4.088***	4.292***	4.312***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
R^2	0.631	0.631	0.661	0.661	0.661	0.688	0.688
Obs	5615	5584	4449	4449	4359	4155	4155

Note. The dependent variable is INTENSITY. All models contain country and year dummies. Estimation is by maximum likelihood in an ordered logit model. p-values based on robust standard errors clustered at the country level are in parentheses. $^{*}p < 10$, $^{**}p < .05$, $^{***}p < .01$.

4.1. **Further robustness tests.** Appendix A presents additional analyses to probe the robustness of these results. Results are summarized in the following tables:

- Table A.1 uses four alternative dependen variables: BATTLE DEATHS (BEST), BATTLE DEATHS, (LOW), INCIDENCE and CONFLICT-SHARE. The first two variables measure the log of the number of battle related deaths, according to the best and low estimates, respectively (Lacina and Gleditsch 2005), whereas INCIDENCE is a dummy that is equal to one whenever there is conflict. CONFLICT-SHARE is the share of years over the period 1992–2010 a group has been involved in conflict.

- Table A.2 re-estimates the models in Table 2 but using G^{I}, the group Gini variable adjustment with the intercept approach, in place of G^{R}. Model 8 in the table uses G^{U}, the unadjusted survey-based group Gini.
- Table A.3 replicates Table 2 but creating the survey measures of G^{R} in a different way. The measures in Table 2 are based on the average of the various surveys for a group applied to all years in order to maximize available data. To avoid the possibility that this measure of group Gini is affected by civil wars that may have occurred prior to the dates of a survey, we have re-computed group inequality so that in year t we take the mean of all available surveys up to year t (and is set to missing if no survey is available prior to year t), yielding a variable we call G^{R}-PRE. Since inequality changes slowly, the correlation between G^{R}-PRE and G^{R}-PRE is very high (r=.94).
- Table A.4 considers different measures of horizontal inequality and group GDP.
- Table A.5 considers a range of alternative control variables and model specifications. Model 1 includes a country-level measure of oil resources, the net value of exports of oil and gas (Ross 2013). Column 2 introduces a variable capturing the number of political transitions in the last 5 years (Cheibubu, Gandhi and Vreeland, 2010). Column 3 adds the distance from the ethnic homeland to the country's capital. Column 4 adds POVERTY2, a country-level measure of poverty (from PovcalNet, World Bank), while column 5 introduces INFANT MORTALITY (from Prio-Grid), which measures group-level infant mortality rate. Column 6 introduces all the previous variables in the model. Finally, column 7 considers an specification with country-year dummies (which is estimated by OLS as the optimization algorithm of the ordered logit model fails to converge in that case).
- Finally, we re-estimated Model 7 in Table 2, dropping each country to assess whether the results are being driven by particular countries.

This battery of robustness tests yields a consistently positive and precisely estimated association between within-group inequality and the intensity or incidence of civil conflict.

5. CORRELATION VERSUS CAUSALITY

While the results presented in the previous section provide robust support for the ER theory, they obviously do not imply causation. As usual, the two central concerns are omitted variable bias and reverse causality. We discuss the two in turn.

5.1. **Omitted variable bias I: Variables related to group inequality.**

We face the standard concern that omitted variables may be affecting both the heterogeneity of incomes within a group and its propensity to be involved in conflict. The identifying assumption – i.e., that within-group inequality is close to random conditional on observable characteristics – does not hold if there is a systematic relationship between group Ginis and other unobserved country or group characteristics. Including country-fixed effects in all specifications partly attenuates the problem by making it possible to rely exclusively on within-country variation to identify the parameters. The specifications also include standard group-level control variables that have been central in previous research on conflict. Nonetheless, omitted group-level variables could still be biasing the results, and this section probes this possibility in two ways.

The first is to examine more carefully the variables that are associated with group-level inequality. The origins of within-group inequality is an issue that has received virtually no theoretical or empirical attention, and it is beyond the scope of this paper to address this issue convincingly. But since the previous models include only variables that previous research has linked to conflict, it is useful to ask whether there are other group-level variables that should (a) have a systematic relationship with group-level Ginis, and (b) could plausibly influence conflict intensity via a pathway other than that posited in ER's theory of labor and capital. If we identify variables that satisfy both these conditions, we must be concerned that these variables could be driving the results in the previous regressions.

Appendix A.2 addresses this issue. The analysis examines the association between a wide range of group-level variables that could plausibly be related at once to within-group inequality and conflict. The analysis shows that (i) a large fraction of the variance of G^R (around 75%) is explained in those regressions, which suggests that we are able to pin down the major sources of variation of G^R, and (ii) among the new variables considered, only one – linguistic heterogeneity within a group – is found to have a precisely estimated relationship with group-level inequality. When linguistic heterogeneity is introduced into the conflict regressions, its coefficient is insignificant and the results for within-group inequality remain robust. This analysis therefore should diminish concerns about the results being driven by omitted variables.

5.2. **Omitted variable bias II: A direct test.** Although we have found that the results for group inequality are robust to including a large number of group and country-level controls, the possibility of omitted variable bias always remains. An alternative way to investigate its possible importance is to look at the stability of the coefficients of interest after the inclusion of additional explanatory variables. To this end, we employ a method recently developed by Oster (2016), which builds on the work by Altonji, Elder and Taber (2005). The method is based on the fact that omitted variable bias is proportional to coefficient movements computed in models with and without controls, scaled by the change in R-squared when controls are included. The idea is that if including controls substantially attenuates the coefficient estimates for WGI, then it is possible that inclusion of more controls would reduce the estimated effect even further. But if the inclusion of controls does not affect the magnitude of coefficient estimates, we can be more confident in suggesting a causal interpretation for the estimated relationship. As emphasized by Oster (2016), scaling by R-squared movements is key to diagnosing the quality of the added controls because adding an irrelevant control might not have an effect on the coefficient of interest, making that coefficient look stable despite the fact that the omitted variable bias can be large. Thus, by scaling by movements in R^2's, it's possible to take into account the quality of the controls. This analysis, found in Appendix A.2.2, suggests it is very unlikely that omitted variables are driving the result.

5.3. **Reverse causality.** Social conflict can disrupt the economy and affect the distribution of income within groups through its effects on wages and employment, destruction of infrastructure, or the confiscation of assets, among other things. But while conflict can clearly affect inequality, it is unclear whether or why this effect should be systematic across groups. A civil war, for example, may hurt the richest within groups (lowering intra-group inequality) or hurt the poorest (increasing intra-group inequality). Though we cannot offer a full theoretical treatment of this issue, we can examine whether civil war influences group-level Ginis in our data because for some countries we have data on inequality at different points in time. Thus, we can measure changes in inequality within groups over time to see if these changes differ with the incidence of civil conflict.

We therefore created a data set where the unit of observation is a country-group, and for each group, we have measured how within-group inequality changes (using both G^R and G^I) between the first and last year for which we have data. This "change in group Gini" variable is the dependent variable that we regress on the proportion of years in which there was civil conflict after the first year for which we have data for a group. The regressions include two control variables: one for the last year for which we have data, and another for the total number of years between our first and last survey for the group. We estimate the model with and without country dummy variables, and with and without groups that were in conflict for the first year we have data. In none of the eight models is the coefficient for the proportion of years in conflict at all precisely estimated, and it switches signs across various models. Although based on a relatively small amount of data, this analysis suggests there is no evidence that past civil conflict has a systematic positive or negative relationship with the level of inequality within a group. Further details are in Appendix A.2.3.

6. Empirical analysis of group inequality and the onset of conflict

The analysis above examined the intensity and incidence of civil conflict, which the ER theory suggests should be linked to within-group inequality. This section turns to the *onset* of conflict, which has been central to studies emphasizing horizontal inequality. Before turning to the role of horizontal inequality, however, it is instructive to consider the relationship between within-group inequality and conflict onset. Esteban and Ray's argument assumes civil conflict exists and links WGI to its *intensity*. The decision to start a new conflict, however, is different from that of continuing or investing more in an existing one, making the relationship between WGI and conflict onset ambiguous. On the one hand, a large degree of within-group inequality might lower the opportunity costs of starting a new conflict. On the other hand, it may deter conflict if actors foresee that fighting has the potential to become very intense and destructive, prompting negotiation to prevent it. Therefore, before turning to what has been the central focus of group-level studies of inequality and conflict onset – i.e., horizontal inequality – we examine whether there is an association between WGI and conflict onset.

6.1. **WGI and conflict onset.** The dependent variable in the onset regressions, also taken from EPR, is ONSET, which equals 1 during the first year in which a group is involved in an armed conflict that results in more than 25 battle-related deaths.[26] Column 1 in Table 3 is similar to column 1 in Table 2, with only country and year indicators as controls, and with the lagged dependent variable replaced by PEACEYEARS, a variable that measures number of years since the last conflict observation. The coefficient of G^R is positive but insignificant and its magnitude is around half of that reported in Table 2. Columns 2 and 3 add all the group-level and country-level controls used in columns 4 to 6 of Table 2. Introducing controls reduces even further the magnitude of the coefficient of G^R, which remains imprecisely measured. Column 4 removes country GDP from the specification in column 3, while column 5 replaces HI(LN) by LINEQ2, a previously used measure of horizontal inequality discussed in Appendix B. Column 6 uses G^I in place of G^R. In none of these regressions are the inequality variables significantly related to conflict onset. Finally, column 7 introduces three additional controls that are frequently employed in onset regressions (see for instance Cederman et al., 2011 and Kuhn and Weidmann 2015): N.EXCL.GROUPS is the number of excluded groups in the country, while POW.BALANCE and POW.BALANCE^2 is the group's demographic power balance with the ethnic group(s) in power as its share of the dyadic population and its square, respectively. While these controls are significant, conclusions about the inequality variables remain unchanged.

The results suggest a very weak relationship between WGI and conflict onset: the parameter estimates for the WGI variables are always positive but their magnitude is around half of that

[26]This variable is called ONSET_DO_FLAG in the EPR data set, and it is missing in years where a group is involved in conflict during the years immediately following the year when conflict is initiated. Thus, the data include group-years where a group either begins involvement in conflict or is not involved in conflict. See Wucherpfennig et al. 2012 for details on conflict measures.

Table 3. Within-group inequality (measured with surveys) and conflict onset

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
G^R	3.256	1.808	2.747	3.065	3.895		2.969
	(0.354)	(0.686)	(0.589)	(0.336)	(0.224)		(0.381)
G^I						4.877	
						(0.239)	
HI(LN)		0.304	0.355	0.353		0.339	0.575
		(0.422)	(0.273)	(0.261)		(0.272)	(0.179)
LINEQ2					0.955		
					(0.363)		
GROUP SIZE		-0.459	0.340	0.293	-0.935	0.162	0.082
		(0.890)	(0.892)	(0.916)	(0.757)	(0.955)	(0.991)
EXCLUDED		1.956***	1.921**	1.934**	1.736**	1.910**	3.899***
		(0.007)	(0.020)	(0.020)	(0.037)	(0.018)	(0.004)
GROUP ELEV. (SD)		0.001	0.001	0.001	0.000	0.001	0.001
		(0.872)	(0.787)	(0.790)	(0.931)	(0.768)	(0.740)
GROUP DIAMONDS		0.729	0.767	0.787	0.678	0.862	2.118
		(0.568)	(0.574)	(0.557)	(0.573)	(0.515)	(0.123)
GROUP OIL		0.212	-0.712	-0.713	-0.794	-0.697	-0.468
		(0.858)	(0.577)	(0.579)	(0.542)	(0.595)	(0.745)
POP			-8.221	-8.235	-7.775	-8.474	-15.190
			(0.367)	(0.366)	(0.363)	(0.347)	(0.176)
GROUP GDP		-0.672	-0.122				
		(0.589)	(0.914)				
GDP			-0.989	-1.114	-0.755	-1.028	-0.024
			(0.829)	(0.791)	(0.862)	(0.805)	(0.996)
XPOLITY			-0.140	-0.140	-0.136	-0.141	-0.269
			(0.419)	(0.419)	(0.439)	(0.419)	(0.140)
POW. BALANCE							17.605*
							(0.059)
POW. BALANCE^2							-23.013**
							(0.035)
N. EXCL. GROUPS							-1.895***
							(0.005)
PEACEYRS	-0.067***	-0.056***	-0.059**	-0.059***	-0.059***	-0.059***	-0.071**
	(0.000)	(0.009)	(0.013)	(0.009)	(0.009)	(0.009)	(0.022)
c	-5.762***	-1.922	9.996	9.969	8.996	8.277	7.922
	(0.002)	(0.873)	(0.779)	(0.780)	(0.803)	(0.814)	(0.850)
R^2	0.200	0.277	0.282	0.282	0.277	0.282	0.361
Obs	1111	844	791	791	791	791	791

Note. The dependent variable is ONSET. All models contain country and year dummies. Estimation is by maximum likelihood using logit. p-values based on robust standard errors clustered at the country level are in parentheses. $^*p < 10$, $^{**}p < .05$, $^{***}p < .01$.

obtained in the incidence regressions and always estimated with substantial error.[27] We also estimated these 7 models using G^R-PRE as the measure of within-group inequality. Although one

[27]We have formally tested whether the coefficients for G^R in the onset and the intensity regressions in Tables 2 and 3 are statistically different. Since the number of observations in the onset regressions is considerably smaller, power is likely to be low. Nonetheless, tests comparing the estimates of G^R in columns 3 and 4 from Table 3 to those obtained

should bear in mind that the remaining data for onset regressions is small, we again find no association between WGI and conflict.

These results contrast sharply with Kuhn and Weidmann's (2015) finding of a positive and significant relationship between conflict onset and their measures of WGI based on nightlight emissions data. We argue in Appendix B.5 that the WGI measures based on geo-coded data suffer serious flaws, which could explain these differences. But the differing results could be also be due to a number of other factors, including that the survey data has fewer observations and that the Kuhn and Weidmann (2015) analysis uses a different set of control variables. It is therefore worth revisiting their analysis, which we do in Appendix A.3. We find that the results from their paper are not robust. Application of Oster's technique suggests that the results are likely to be non-robust due to omitted variable bias. Indeed, we find that when our baseline controls are added to their regressions, the association disappears. In addition, results also vanish when one particular group is dropped from the sample. Thus, we find very consistent results, using both nighlight and survey-based WGI estimates, suggesting that there is no association between WGI and the onset of ethnic conflict.

6.2. **Horizontal inequality and conflict onset.** The positive relationship between horizontal inequality and conflict onset has found substantial support in group-level empirical tests relying on EPR group-level data (see CWG, Cederman, Weidmann and Bormann 2015 ("CWB"), KW, and Cederman, Gleditsch and Buhaug 2013 (CGB). However, other research summarized in Section 2.2 questions this relationship from both a theoretical and an empirical perspective.[28] The findings in Tables 3 and A.10 (in Appendix A.3) are consistent with the latter literature: the coefficients for horizontal inequality vary widely in size and in sign and are never significant, raising questions about the robustness of previous findings. This section therefore offers additional tests of HI arguments by using the most recent version of the EPR data set. Details regarding the data and

in the analogous regressions in Table 2 (columns 6 and 7) deliver p-values of .026 and .06, respectively, depending on the specification employed, which suggests there is reasonable evidence in the data to reject equality of coefficients.

[28]See Esteban and Ray (2011), Mitra and Ray (2014) and Adhvaryu et al. (2017).

Table 4. Horizontal inequality and conflict onset using EPR data

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
HI(LN)	0.052							
	(0.714)							
LINEQ2		-0.619***						
		(0.000)						
LOW			-0.389***					
			(0.004)					
HIGH			-0.281					
			(0.220)					
$\overline{\text{LOW}}$				-0.658***				
				(0.003)				
$\overline{\text{HIGH}}$				-0.157				
				(0.524)				
HI(LN), SMALL					0.056			
					(0.782)			
LINEQ2, SMALL						-1.375		
						(0.177)		
LOW, SMALL							-0.721	
							(0.127)	
HIGH, SMALL							-0.472	
							(0.560)	
$\overline{\text{LOW}}$, SMALL								-0.714*
								(0.057)
$\overline{\text{HIGH}}$,SMALL								-0.370
								(0.342)
EXCLUDED	0.977**	1.004**	0.994**	0.981**	0.761*	0.738	0.740	0.762*
	(0.017)	(0.016)	(0.017)	(0.018)	(0.081)	(0.110)	(0.107)	(0.099)
DOWNGRADED	1.976***	2.048***	2.055***	2.065***	1.912***	1.979***	1.984***	2.136***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
GROUP SIZE	-1.185	-1.288	-1.242	-1.174	-2.416*	-2.617**	-2.526*	-2.020
	(0.318)	(0.297)	(0.310)	(0.355)	(0.050)	(0.039)	(0.060)	(0.217)
POSTWAR	-0.064	-0.025	-0.027	-0.066	0.065	0.029	0.029	0.048
	(0.678)	(0.881)	(0.878)	(0.678)	(0.742)	(0.877)	(0.877)	(0.789)
ONGOING CONFLICT	-0.518	-0.520	-0.520	-0.527	-0.446	-0.445	-0.445	-0.452
	(0.283)	(0.286)	(0.287)	(0.281)	(0.362)	(0.361)	(0.359)	(0.342)
GDP	-0.342	-0.160	-0.175	-0.252	-1.012	-0.845	-0.858	-0.911
	(0.473)	(0.733)	(0.705)	(0.581)	(0.129)	(0.165)	(0.160)	(0.128)
POP	-1.986	-2.064	-2.063	-1.948	-2.181	-2.153	-2.130	-2.128
	(0.176)	(0.158)	(0.160)	(0.189)	(0.180)	(0.185)	(0.189)	(0.189)
PEACEYEARS	-0.176***	-0.174***	-0.174***	-0.175***	-0.149***	-0.153***	-0.154***	-0.155***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
c	10.776	9.597	10.463	10.848	17.477*	16.244*	17.481*	17.720*
	(0.177)	(0.213)	(0.173)	(0.152)	(0.075)	(0.079)	(0.067)	(0.055)
R^2	0.264	0.269	0.269	0.270	0.236	0.239	0.239	0.242
Obs	4721	4721	4718	4708	2635	2635	2635	2627

Note. The dependent variable is ONSET. All models contain country fixed effects and a natural cubic spline with three knots. Estimation is by maximum likelihood in a logit model. All the inequality measures have been computed using data from GECON (obtained from `https://growup.ethz.ch`. HI measures ending by "small" are computed by dropping groups whose population is smaller than less than 500.000 people. p-values based on robust standard errors clustered at the country level are in parentheses. $^{*}p < 10$, $^{**}p < .05$, $^{***}p < .01$.

variable construction (and how they differ from CWB and other papers using EPR data) are found in Appendix B.6.[29]

Table 4 presents models that include the same set of controls used in CWB, although the models differ from CWB by including country fixed effects and by lagging right-hand side variables where possible. Models 1-4 are identical to each other except that each uses a different measure of horizontal inequality: model 1 uses HI(LN) (our preferred measure, discussed above); model 2 uses LINEQ2 (the measure advocated in CWG and KW); model 3 uses LOW and HIGH (the measures advocated in CGB and CWB), which are defined as LOW $= max\{1, \frac{G}{G_g}\}$, and HIGH $= max\{1, \frac{G_g}{G}\}$, where G_g denotes group's g GDP per capita and G is the (unweighted) average of the per capita GDP of all groups; model 4 uses the corrected measures of $\overline{\text{LOW}}$ and $\overline{\text{HIGH}}$.[30] The results differ substantially from previous research. In no model is the coefficient for a horizontal inequality variable both positive and precisely estimated, and in 3 of the 4 models there is a coefficient for a horizontal inequality measure that is *negative* and precisely estimated. Since the coefficient of LOW is negative and significant in models 3 and 4, the results here suggest a very different relationship between relative poverty and conflict than the one suggested by the grievance arguments: very poor groups might enter conflict less because they simply lack the means to oppose the state.

Models 1-4 include all available groups. CWB, however, argue that the EPR measures of HI likely suffer measurement error for small groups and exclude such groups from the analysis (see also CWG). Models 5-8 therefore re-estimate models 1-4, but excluding small groups. The coefficients for the HI variables are not positive and precisely estimated in any model.[31] A number of additional tests produce similar results. First, CWB use a former version data that ends in 2009, whereas our EPR data extend through 2013. We therefore estimated the models in Table 4 using

[29]As described in detail in Appendix B.6, there are three main differences between our HI measures and those in CWB: 1) we use a more recent version of the EPR data (2014 update 2); 2) in regressions where small groups are dropped, our HI measures are created without those groups and 3) we exploit the time variation of the GECON GDP data and compute time-varying HI measures (that are lagged in the regressions).

[30]As described in Appendix B.3, the "corrected" measures differ from the original ones in the definition of the average GDP to which the GDP of group g is compared to. See Appendix B.3 for details.

[31]We also used the nightlights data to create the four different measures of Low and High and in Table 4 (from models 3, 4, 7 and 8). When these variables are used, the results are similarly unsupportive of the horizontal inequality arguments. Results can be found in the replication files.

only the data up through 2009. Second, CWG uses a slightly different set of controls than CWB. We therefore estimated the models in Table 4 by including (a) any control variable that appears in CWB or CWG, and (b) only the controls from CWB. Across all 36 models we estimate, we never find a coefficient for a horizontal inequality variable that is positive and precisely estimated.[32] Our analysis therefore raises the clear possibility that in the study of conflict onset, the empirical support for horizontal inequality arguments found in previous studies using EPR data is not robust. Further research could explore if this lack of robustness is due to measurement error or to the fact that the sample period considered in the empirical analysis is not long enough. However, we believe that the problem is more fundamental and that the relationship between differences in income across groups and conflict onset is much more subtle than previously thought, as recent research has suggested (Esteban and Ray 2011, Adhvaryu et al. 2017).

7. Conclusion

Group level studies of ethnic civil conflict have typically focused on the role of economic grievance in sparking civil wars. Our analysis, however, reveals no robust support for arguments that inequality across groups is linked to the start of civil conflict. Civil wars, however, vary a great deal in their severity, making it crucial to also understand group-level factors that influence the intensity of civil wars, whatever the reason such wars begin. To this end, our analysis reveals strong support for Esteban and Ray's (2011) argument that unequal groups should be involved in the most severe civil conflicts. We hope these findings motivate additional studies aimed at improving our understanding of how the capacity of groups to fight is related to the level of death and destruction from civil conflicts.

[32] In addition to using HI variables obtained from three different data sources (G-Econ, nightlight emissions and survey data), CWB construct new proxies by computing weighted averages of (a function of) Low/High obtained from those datasets. However, notice that Low/High are very nonlinear measures (recall that LOW $= \max\{1, G/G_g\}$ and HIGH $= \max\{1, G_g/G\}$). Given these non-linearities, averaging first the ratios of GDPs and then computing the Low/High measures or viceversa (as CWB do), can lead to very different conclusions. For instance, CWB's approach implies that if one group is poorer that the average according to one dataset but richer than the average according to another, then the weighted average of these measures will result in Low/High measures that are *both* larger than 1, i.e., that group will be classified simultaneously as relatively poor *and* as relatively rich. These problems will be solved if the ratios of GDPs are averaged first and then HI measures based on these ratios are computed. Given the problems inherent in CWB combined measures, we prefer to use HI variables obtained from a unique dataset.

8. References

Acemoglu, Daron, and James Robison. 2005. "Economic Origins of Dictatorship and Democracy." *Cambridge University Press.*

Adhvaryu, Achyuta, James Fenske, Gaurav Khannax and Anant Nyshadham 2017. "Resources, Conflict, and Economic Development in Africa." Mimeo.

Allison, Graham. 2017. "Can America and China Escape Thucydides's Trap?" *Houghton Mifflin Harcourt.*

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2005. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy*, 113(1): 151-184.

Angel, Shlomo. 2012. "Planet of Cities." *Lincoln Institute of Land Policy, MA: Cambridge.*

Angrist, Joshua David, and Jörn-Steffen Pischke. 2009. "Mostly harmless econometrics: an empiricist's companion." *Princeton University Press.*

Arbatli, Cemal E., Quamrul Ashraf and Oded Galor. 2013. "The Nature of Civil Conflict." Mimeo.

Bazzi, Samuel, and Christopher Blattman. 2014. "Economic Shocks and Conflict: Evidence from Commodity Prices." *American Economic Journal: Macroeconomics*, 6(4): 1-38.

Bellows, John, and Edward Miguel. 2009. "War and Local Collective Action in Sierra Leone." *Journal of Public Economics*, 93 (11-12): 1144-1157.

Brubaker, Rogers and David D. Laitin. 1998. "Ethnic and Nationalist Violence." *Annual Review of Sociology*, 24, 42-452.

Cederman, Lars-Erik, Gleditsch Kristian S., and Halvard Buhaug. 2013. "Inequality, Gruevances, and Civil War." *Cambridge University Press.*

Cederman, Lars-Erik, Brian Min, and Andreas Wimmer. 2009. Ethnic Power Relations dataset, hdl:1902.1/11796.

Cederman, Lars-Erik, Nils B. Weidmann, and Nils-Christian Bormann. 2015. "Triangulating Horizontal Inequality: Toward Improved Conflict Analysis." *Journal of Peace Research*, 52(6): 806:21.

Cederman, Lars-Erik, Nils B. Weidmann, and Kristian S. Gleditsch 2011. "Horizontal Inequalities and Ethnonationalist Civil War: A Global Comparison." *American Political Science Review*, 105(3) 478-495.

Cheibub, Jose Antonio, Jennifer Gandhi and James R. Vreeland 2010. "Democracy and dictatorship revisited" *Public Choice*, 143, Issue 1?2, 67–101

Cramer, Christopher. 2003. "Does Inequality Cause Conflict?" *Journal of International Development*, 15: 397-412.

Collier, Paul, and Anke Hoeffler. 2004. "Greed and Grievance in Civil War." *Oxford Economics Papers*, 56(4): 563-595.

Cunningham, David E., Kristian Skrede Gleditsch and Idean Salehyan. 2009. "A Dyadic Analysis of Civil War Duration and Outcome." *Journal of Conflict Resolution*, 53(4): 570-597.

Dahrendorf, Ralf. 1959. "Class and class conflict in industrial society." *Stanford, CA: Stanford University Press.*

Dal Bó, Ernesto, and Dal Bó Pedro. 2011. "Workers, Warriors, And Criminals: Social Conflict In General Equilibrium." *Journal of European Economic Association*, 4(8): 646-677.

Deininger, Klaus, and Lyn Squire. 1996. "A New Data Set Measuring Income Inequality." *The World Bank Economic Review*, 10(3): 565-591.

Dube, Oeindrila, and Juan Vargas. 2013. "Commodity Price Shocks and Civil Conflict: Evidence from Colombia." *The Review of Economic Studies*, 80(4): 1384-1421.

Esteban, Joan, and Debraj Ray. 2011. "A Model of Ethnic Conflict." *Journal of the European Economic Association*, 9(3):496-521.

Esteban, Joan, Laura Mayoral, and Debraj Ray. 2012. " Ethnicity and Conflict: An empirical Study." *American Economic Review*, 102(4): 1310-42.

Fearon, James D. 2003. "Ethnic and Cultural Diversity by Country." *Journal of Economic Growth*, 8(2): 195–222.

Fearon, James D. 2004. "Why do some civil wars last so much longer than others?" *Journal of Peace Research* 41(3): 275-302.

Fearon, James D., and David D. Laitin. 2000. "Violence and the Social Construction of Ethnic Identity." *International Organization*, Cambridge University Press, vol. 54(04): 845-877.

Fearon, James D., and David D. Laitin. 2003. "Ethnicity, Insurgency, and Civil War." *American Political Science Review*, 97(1): 75-90.

Filmer, Deon, and Lant H. Pritchett. 2001. "Estimating Wealth Effects Without Expenditure Data-Or Tears: An Application to Educational Enrollments in States of India." *Demography*, 38(1):115-132.

Fjelde, Hanne, and Østby Gudrun. 2014. "Socioeconomic Inequality and Communal Conflict: A Disaggregated Analysis of Sub-Saharan Africa, 19902008." *International Interactions*, 40(5):737-762.

Gleditsch, Nils P., Peter Wallensteen, Mikael Eriksson, Margareta Sollenber, and Håvard Strand. 2002. "Armed Conflict 1946-2001: A New data set." *Journal of Peace Research*, 39(5): 615-637 (accessed October 1, 2010).

Gurr, Ted R. 1970. "Why Men Rebel." *Princeton: Princeton University Press.*

Gurr, Ted R. 1980. "Why Men Rebel." *Handbook of Political Conflict: Theory and Research*, New York Free Press.

Huber, John D. 2017. "Exclusion by Elections: Inequality, Ethnic Identity and Democracy." *New York: Cambridge University Press.*

Humphreys, Macarta and Jeremy M. Weinstein. 2008. "Who Fights? The Determinants of Participation in Civil War." *American Journal of Political Science*, 52, 436-455.

Justino, Patricia. 2009. "Poverty and Violent Conflict: A Micro-Level Perspective on the Causes and Duration of Warfare." *Journal of Peace Research, Peace Research Institute Oslo*, 46(3):315-333.

Kuhn, Patrick M., and Nils B. Weidmann. 2015. "Unequal we fight: Between-and within-group inequality and ethnic civil war." *Political Science Research & Methods*, 3(3): 543-568.

Lacina, Bethany. 2006. "Explaining the Severity of Civil Wars." *Journal of Conflict Resolution*, 50(2): 276-289.

Lacina, Bethany and Nils P., Gleditsch. 2005. "Monitoring Trends in Global Combat: A New Dataset of Battle Deaths." *European Journal of Population / Revue européenne de Démographie*, 21(2): 145-166.

Lichbach, Mark I. 1989. "An Evaluation of 'Does Economic Inequality Breed Political Conflict?' Studies." *World Politics*, 41(4): 431-470.

McKenzie, David J. 2005. "Measuring Inequality with Asset Indicators." *Journal of Population Economics*, 18(2):229-260.

Melvern, Linda. 2000. "A people betrayed: the role of the West in Rwanda's genocide." *New York, N.Y.: Zed Books.*

Morelli, Massimo and Dominic Rohner. 2015. "Resource Concentration and Civil Wars." *Journal of Development Economics*, 117(C): 32-47.

Nordhaus, William D. 2006. "Geography and macroeconomics: New data and new findings." *Proceedings of the National Academy of Sciences of the USA*, 103(10): 3510-3517.

Østby, Gudrun. 2008. "Polarization, Horizontal Inequalities and Violent Civil Conflict." *Journal of Peace Research*, 45(2): 143-162.

Østby, Gudrun, Ragnhild Nordas, and Jan Ketil Rod. 2009. "Regional Inequalities and Civil Conflict in Sub-Saharan Africa." *International Studies Quarterly*, 53(2): 301-324.

Oster, Emily. 2016. "Unobservable Selection and Coefficient Stability: Theory and Evidence." *Journal of Business Economics and Statistics,* Forthcoming.

Penn World Table. 2011. Dataset,

`https://pwt.sas.upenn.edu/php_site/pwt_index.php`.

Polity IV. "Polity IV Project: Political Regime Characteristics and Transitions, 1800-2009, " (accessed October 1, 2011).

`http://www.systemicpeace.org/polity/polity4.htm`

Ross, Michael and Paasha Mahdavi. 2015, "Oil and Gas Data, 1932-2014", doi:10.7910/DVN/ZTPW0Y, Harvard Dataverse, V2.

Solt, Frederik. 2009. "Standardizing the World Income Inequality Database." *Social Science Quarterly*, 90(2): 231–242.

Stewart, Frances. 2000. "Dynamic Interactions between the Macro-Environment, Development Thinking and Group Behavior." *Development Working Papers*, 143, University of Milano.

Stewart, Frances. 2002. "Horizontal Inequalities: A Neglected Dimension of Development." *Annual Lecture No. 5, UNU World Institute for Development Economics Research.*

Verwimp, Philip. 2005. "An economic profile of peasant perpetrators of genocide. Micro-level evidence from Rwanda." *Journal of Development Economics*, 77(2): 297-323.

Vreeland, James R. 2008. "The Effect of Political Regime on Civil War". In *Journal of Conflict Resolution*, 52(3): 401-425.

Weidmann, Nils B., Jan Ketil Rød, and Lars-Erik Cederman. 2010. "Representing Ethnic Groups in Space: A New Dataset." *Journal of Peace Research*, 47(4): 491-99.

Wimmer, Andreas, Lars-Erik Cederman and Brian Min. 2009. "Ethnic politics and armed conflict. A configurational analysis of a new global dataset." *American Sociological Review*, 74(2):316-337.

Wimmer, Andreas. 2015. "Race-centrism: a critique and a research agenda." *Ethnic and Racial Studies*, 38(13):2186-2205.

Wintrobe, Ronald. 1995. "Some Economics of Ethnic Capital Formation and Conflict." *Nationalism and Rationality*, edited by Albert Breton, Gianluigi Galeotti, and Ronald Wintrobe: 43-70.

World Bank. 2013. Dataset,
`http://iresearch.worldbank.org/PovcalNet/index.htm?0,2`.

Wooldridge, Jeffrey M. 2002. "Econometric Analysis of Cross Section and Panel Data (Second Edition)." *MIT press.*

J Wucherpfennig, NB Weidmann, L Girardin, LE Cederman, A Wimmer. 2011. "Politically Relevant Ethnic Groups Across Space and Time: Introducing the GeoEPR Dataset". *Conflict Management and Peace Science* , 28 (5), 423-437.

Yanagizawa-Drott, David. 2012. "Propaganda and conflict: Theory and Evidence from the Rwandan Genocide." *Harvard University.*

APPENDIX A. GROUP INEQUALITY AND CONFLICT: ADDITIONAL ANALYSIS

This appendix provides the additional empirical analyses discussed in the text. Section A.1 contains the robustness tests for the conflict incidence and intensity regressions discussed in Section 4 of the main text; Section A.2 examines the issues of correlation vs causation that are discussed in the paper's Section 5. And section A.3 provide additional analysis of WGI and conflict onset.

A.1. **Robustness checks.** As a first step in probing the robustness of results in Section 4, Table A.1 considers alternative dependent variables. Columns 1–4 consider a more direct measure of intensity, (the log of) the number of battle related deaths (Lacina and Gleditsch, 2005). As there is considerable uncertainty in the measurement of this variable, Lacina and Gleditsch (2005) provide three different estimates: a lower bound, an upper bound and the "best estimate" (that is very often missing).[33] Columns 1–2 employ the "best" estimate (and whenever missing, the "low" estimate is used instead), while columns 3–4 employ the "low estimate." Columns 5–7 use a binary variable capturing conflict incidence, INCIDENCE, which takes the value 1 if the group is involved in conflict against the state resulting in more than 25 battle related deaths. Finally, columns 7 and 8 collapse the time dimension of the data and consider as dependent variable the share of years over the period 1992–2010 a group has been involved in conflict.[34] Estimation is carried out by OLS (columns 1–4 and 7–8) and maximum likelihood in a logit model (columns 5–6). Our conclusions remain robust to using alternative dependent variables capturing the intensity and the incidence of conflict. Finally, Figure A.1 provides a graphical representation of the relation between CONFLICT-SHARE and G^R.

Tables A.2 and A.3 reestimate Table 2 using alternative measures of WGI. Models 1-7 in Table A.2 use G^I, the group Gini variable adjusted with the intercept approach. Model 8 uses G^U, the

[33]For relatively low intensity conflicts with uncertain casualty counts, PRIO assigns a death count of exacltly 25, and 19% of all conflict observations in the battle deaths data have exactly 25 battle deaths, meaning that in 19% of the lower intensity cases, precise counts of deaths were unavailable. Similarly, for higher intensity conflicts, 8% of observations have assigned exactly 1000 battle deaths to a conflict observation. For this reason, the "discrete" version of the intensity variable used above has been adopted in the literature.

[34]More specifically, CONFLICT-SHARE is the average of INCIDENCE over the period 1992–2010 and is employed in regressions where the unit of analysis is the group. All time-varying variables are considered at the beginning of the sample.

Table A.1. Within group inequality, number of battle deaths and conflict incidence

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
G^R	0.460*	0.604*	0.436*	0.548*	5.967**	5.610**	0.206*	0.287**
	(0.081)	(0.051)	(0.078)	(0.057)	(0.026)	(0.029)	(0.054)	(0.049)
HI(LN)	0.013	0.015	0.011	0.012	0.141	0.064	0.007	0.007
	(0.204)	(0.224)	(0.239)	(0.284)	(0.414)	(0.800)	(0.243)	(0.398)
EXCLUDED GROUP		0.088		0.078		0.684		0.052*
		(0.196)		(0.190)		(0.342)		(0.098)
GROUP SIZE		-0.040		-0.037		-0.877		0.004
		(0.581)		(0.588)		(0.719)		(0.949)
GROUP ELEV. (SD)		-0.000		-0.000		-0.001		-0.000
		(0.405)		(0.381)		(0.750)		(0.704)
GROUP DIAMONDS		0.032		0.030		0.819		0.006
		(0.420)		(0.419)		(0.436)		(0.863)
GROUP OIL		-0.007		-0.005		0.292		0.024
		(0.893)		(0.919)		(0.694)		(0.441)
POP		-0.082		0.010		-1.326		0.140
		(0.541)		(0.903)		(0.542)		(0.641)
XPOLITY		-0.005		-0.006		-0.127*		0.006
		(0.474)		(0.354)		(0.083)		(0.646)
GDP		-0.032		0.021		0.005		0.026
		(0.663)		(0.726)		(0.997)		(0.903)
BATTLE DEATHS (BEST), LAG	0.676***	0.701***						
	(0.000)	(0.000)						
BATTLE DEATHS (LOW), LAG			0.655***	0.694***				
			(0.000)	(0.000)				
INCIDENCE(LAG)					4.579***	4.689***		
					(0.000)	(0.000)		
c	-0.133	-0.025	-0.163	-0.485	-8.870***	-6.254	-0.089	-1.056
	(0.303)	(0.968)	(0.203)	(0.378)	(0.000)	(0.646)	(0.500)	(0.604)
R^2	0.628	0.678	0.621	0.681	0.525	0.583	0.388	0.402
Obs	5556	4145	5556	4145	1731	1410	336	256

Note. The dependent variable is BATTLE DEATHS, (BEST), BATTLE DEATHS, (LOW), INCIDENCE and CONFLICT-SHARE in columns 1–2, 3–4 and 5–7, respectively. All models contain country and year dummies. Estimation of columns 1–4 and 6–8 is by OLS and columns 5–6 by maximum likelihood in a logit model. P-values based on robust standard errors clustered at the country level are in parentheses (except in columns 7 and 8, due to the very small number of observations in those regressions.) $^{*}p < 10$, $^{**}p < .05$, $^{***}p < .01$.

unadjusted survey-based group Gini. The results for WGI are robust to using either measure. Model 8 suggests that the relationship between WGI and conflict incidence exists in the raw data and is not being driven by decisions with respect to adjusting the heterogenous surveys, although the heterogeneity in the surveys obviously warrants making such adjustments.

The measures of WGI used so far are based on the average of the various surveys for a group, and this average is applied to all years, even those that precede the date of the survey. Since

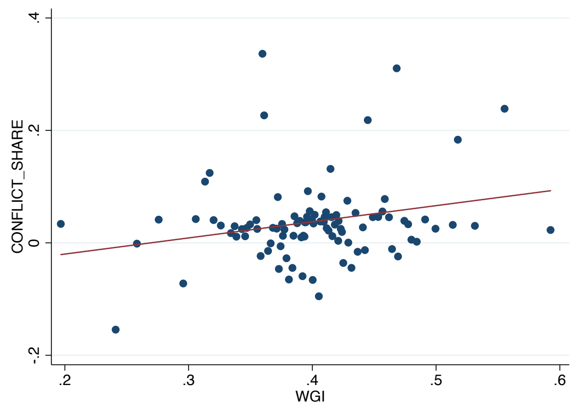


Figure A.1. The relation between CONFLICT-SHARE and G^R. The plot represents the residuals of regressing G^R and CONFLICT-SHARE on country and year fixed effects as well as all the controls in column 7 from Table 2. The dots represent percentiles.

the surveys cover a relatively short period (1992 to 2010) and inequality variables are known to evolve slowly, this approach is a reasonable way to maximize the available data. Nonetheless, reverse causality is a valid concern. To avoid this possibility, we re-computed the inequality variables so that in year t we take the mean of all available surveys in year t and in prior years.[35] Since inequality changes slowly, the correlation between our baseline group Gini and the one just described (denoted as G^R-PRE) is very high (r=.94). Table A.3 replicates Table 2 using these redefined measures of inequality; despite the fact that the sample size reduces considerably, the conclusions regarding within-group inequality are generally robust.

Table A.4 considers different measures of horizontal inequality and GDP. As noted above, GROUP GDP and GDP (a country-level measure) are correlated at .96, and similar results are obtained regardless of whether group GDP or country GDP is included. We can see this in Model 1 of Table A.4, which re-estimates model 7 in Table 2, but substitutes GROUP GDP for GDP. The remaining models in the table re-estimate our baseline specification (model 7 in Table 2) with different measures horizontal inequality : model 2–3 use LINEQ2 (the measure of horizontal inequality in CWG and KW), computed using data either from the surveys (model 2) or GECON (model 3), and models 4–5 consider the LOW and HIGH variables (the measure of horizontal inequality

[35] If no survey is available for a group until time t, observations up to t are set to missing.

Table A.2. Alternative ways of computing Within Group Inequality

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
G^I	7.094**	6.628**	6.664***	6.843**	7.910**	7.969*	6.181**	
	(0.010)	(0.013)	(0.009)	(0.045)	(0.038)	(0.052)	(0.020)	
G^U								7.099***
								(0.004)
HI(LN)		0.099	0.031	0.029	0.061	0.040	0.048	0.040
		(0.550)	(0.869)	(0.875)	(0.748)	(0.851)	(0.831)	(0.860)
GROUP SIZE			-1.605	-1.646	-1.807	-0.883	-0.573	-0.666
			(0.446)	(0.422)	(0.419)	(0.730)	(0.807)	(0.778)
GROUP ELEV. (SD)			-0.001	-0.001	-0.001	-0.000	-0.000	-0.000
			(0.694)	(0.666)	(0.636)	(0.803)	(0.904)	(0.957)
GROUP DIAMONDS			1.158	1.158	1.286	1.093	1.060	0.995
			(0.306)	(0.304)	(0.255)	(0.344)	(0.360)	(0.385)
GROUP OIL			0.389	0.387	0.425	0.266	0.271	0.287
			(0.498)	(0.496)	(0.470)	(0.677)	(0.672)	(0.653)
GROUP GDP				0.061	0.471	0.610		
				(0.922)	(0.500)	(0.401)		
POP					2.879	-1.408	-1.430	-1.515
					(0.451)	(0.559)	(0.548)	(0.522)
GDP					-1.057	0.223	0.825	0.861
					(0.389)	(0.897)	(0.570)	(0.552)
XPOLITY						-0.127*	-0.126*	-0.126*
						(0.054)	(0.055)	(0.055)
EXCLUDED GROUP						0.836	0.804	0.757
						(0.233)	(0.266)	(0.277)
INTENSITY(LAG)	3.954***	3.946***	4.147***	4.148***	4.120***	4.325***	4.335***	4.321***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
R^2	0.630	0.630	0.660	0.660	0.659	0.687	0.686	0.688
Obs	5615	5584	4449	4449	4359	4155	4155	4155

Note. The dependent variable is INTENSITY. All models contain country and year dummies. G^I and G^U denote group Ginis adjusted using the intercept approach and without any adjustment, respectively. Estimation is by maximum likelihood in an ordered logit model. p-values based on robust standard errors clustered at the country level are in parentheses. $^*p < 10$, $^{**}p < .05$, $^{***}p < .01$.

in CGB and CWB), also computed using data from the surveys (column 4) or from GECON (5). Across all five models, the coefficient for WGI remains positive and significant at conventional levels. Similar conclusions are obtained if one controls for group GDP rather than country GDP.

Table A.5 considers a range of alternative variables and model specifications. Column 1 includes a variable measuring the net value of oil and gas exports per capita, OIL/GAS EXPORTS. The data are from Ross and Mahdavi (2015). Column 2 adds POL. TRANSITIONS, the number of political transitions in the previous 5 years (Cheibubu, Gandhi and Vreeland, 2010). Column 3 controls for the the distance from the ethnic homeland to the country capital (DIST.CAP). Model

Table A.3. Estimating models in Table 2 using only past surveys to compute G^R

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
G_t^R-PRE	3.001**	2.922*	4.244*	4.795**	4.414*	3.921*	3.602
	(0.041)	(0.053)	(0.086)	(0.028)	(0.058)	(0.056)	(0.112)
HI_t(LN)-PRE		0.082	0.007	0.040	0.019	-0.028	-0.051
		(0.272)	(0.945)	(0.753)	(0.882)	(0.843)	(0.652)
GROUP SIZE			-2.075	-2.276	-2.294	-0.955	-0.848
			(0.488)	(0.434)	(0.422)	(0.734)	(0.770)
GROUP ELEV. (SD)			-0.002	-0.002	-0.002	-0.001	-0.001
			(0.174)	(0.173)	(0.219)	(0.441)	(0.430)
GROUP DIAMONDS			0.277	0.312	0.283	0.234	0.213
			(0.753)	(0.719)	(0.741)	(0.785)	(0.810)
GROUP OIL			0.327	0.359	0.374	0.152	0.134
			(0.586)	(0.547)	(0.534)	(0.808)	(0.831)
GROUP GDP				0.302	0.170	0.177	
				(0.399)	(0.657)	(0.661)	
POP					5.360	5.436	5.413
					(0.316)	(0.358)	(0.358)
GDP					2.633	2.554	2.713
					(0.257)	(0.325)	(0.276)
XPOLITY						-0.125**	-0.126**
						(0.026)	(0.028)
EXCLUDED GROUP						0.897	0.891
						(0.238)	(0.238)
INTENSITY(LAG)	4.269***	4.262***	4.631***	4.631***	4.656***	4.725***	4.726***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
R^2	0.670	0.671	0.717	0.717	0.718	0.726	0.726
Obs	4027	4027	3178	3178	3137	3012	3012

Note. The dependent variable is INTENSITY. All models contain country and year dummies. Estimation is by maximum likelihood in an ordered logit model. G_t^R-PRE (HI_t(LN)-PRE) is computed such that in year t the mean of all available surveys in year t and in prior years is computed. p-values based on robust standard errors clustered at the country level are in parentheses. $^*p < 10$, $^{**}p < .05$, $^{***}p < .01$.

4 adds POVERTY2, a country-level measure of poverty (from PovcalNet, World Bank), while column 5 introduces INFANT MORTALITY (from Prio-Grid), which measures group-level infant mortality rate. Column 6 includes all the previous variables in the same specification. Model 7 introduces country-year fixed effects and is estimated by OLS, as the optimization algorithm of the ordered logit fails to converge. The results regarding WGI remain robust to the inclusion of these additional controls.

Table A.4. Alternative measures of Group GDP and Horizontal Inequality

	(1)	(2)	(3)	(4)	(5)
G^R	7.790**	5.867**	5.713**	9.870**	3.906*
	(0.019)	(0.019)	(0.018)	(0.028)	(0.071)
GROUP GDP	0.917				
	(0.159)				
GDP		1.014	1.074	0.639	0.739
		(0.485)	(0.409)	(0.661)	(0.613)
HI(LN)	-0.009				
	(0.965)				
LINEQ2		-0.562			
		(0.367)			
LINEQ2 (GECON)			2.049		
			(0.276)		
LOW				-0.775	
				(0.196)	
HIGH				-0.045	
				(0.911)	
LOW (GECON)					0.175
					(0.878)
HIGH (GECON)					-3.437
					(0.177)
GROUP SIZE	-1.145	-0.850	-0.464	-1.674	-1.111
	(0.651)	(0.752)	(0.858)	(0.541)	(0.684)
GROUP ELEV. (SD)	-0.001	-0.001	-0.001	-0.000	-0.001
	(0.554)	(0.522)	(0.352)	(0.754)	(0.160)
GROUP DIAMONDS	0.929	0.872	0.703	0.831	0.681
	(0.397)	(0.434)	(0.499)	(0.451)	(0.425)
GROUP OIL	0.311	0.259	0.119	0.393	0.312
	(0.632)	(0.696)	(0.846)	(0.550)	(0.581)
POP	-1.650	-1.773	-1.656	-2.144	-0.603
	(0.473)	(0.447)	(0.465)	(0.370)	(0.812)
XPOLITY	-0.124*	-0.121*	-0.125*	-0.127**	-0.118*
	(0.061)	(0.061)	(0.050)	(0.050)	(0.076)
EXCLUDED GROUP	0.893	0.891	0.916	0.935	0.938
	(0.194)	(0.213)	(0.162)	(0.194)	(0.149)
INTENSITY(LAG)	4.292***	4.302***	4.245***	4.271***	4.208***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
R^2	0.688	0.688	0.691	0.690	0.695
Obs	4155	4155	4046	4176	4046

Note. The dependent variable is INTENSITY. All models contain country and year dummies. Estimation is by maximum likelihood in an ordered logit model. p-values based on robust standard errors clustered at the country level are in parentheses. $^{*}p < 10$, $^{**}p < .05$, $^{***}p < .01$.

Table A.5. Adding additional control variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
G^R	5.472**	5.139**	4.102**	5.553**	5.775***	3.886*	0.128***
	(0.012)	(0.029)	(0.037)	(0.018)	(0.007)	(0.072)	(0.000)
HI(LN)	0.004	0.016	-0.066	0.008	-0.031	-0.167	0.003
	(0.985)	(0.945)	(0.763)	(0.972)	(0.896)	(0.469)	(0.165)
OIL/GAS EXPORTS	0.001					0.002	
	(0.656)					(0.106)	
POL. TRANSITIONS		0.093				-0.007	
		(0.854)				(0.990)	
DIST. CAP			0.003***			0.003***	
			(0.003)			(0.001)	
POVERTY2				0.040		0.041	
				(0.136)		(0.232)	
INFANT MORTALITY					0.006***	0.008***	
					(0.008)	(0.000)	
GROUP OIL		0.309	0.302	0.302	0.527		-0.004
		(0.644)	(0.642)	(0.647)	(0.380)		(0.612)
EXCLUDED GROUP	0.874	0.825	0.736	0.823	0.417	0.526	0.028***
	(0.204)	(0.248)	(0.354)	(0.248)	(0.502)	(0.418)	(0.000)
GROUP SIZE	-0.686	-0.845	1.465	-0.756	-1.075	1.754	-0.000
	(0.763)	(0.717)	(0.592)	(0.747)	(0.649)	(0.466)	(0.985)
GROUP ELEV. (SD)	-0.001	-0.001	-0.002	-0.000	-0.002	-0.004	-0.000
	(0.757)	(0.737)	(0.229)	(0.772)	(0.330)	(0.148)	(0.378)
GROUP DIAMONDS	0.904	0.895	0.955	0.960	0.548	-0.163	0.012
	(0.404)	(0.441)	(0.329)	(0.397)	(0.648)	(0.856)	(0.105)
POP	-1.567	-3.042	-1.066	-0.515	-1.927	-1.029	
	(0.492)	(0.353)	(0.634)	(0.846)	(0.402)	(0.764)	
XPOLITY	-0.122**	-0.120*	-0.118*	-0.121*	-0.135**	-0.134**	
	(0.047)	(0.090)	(0.064)	(0.051)	(0.032)	(0.011)	
GDP	0.871	0.677	0.440	0.926	0.816	-0.486	
	(0.549)	(0.686)	(0.766)	(0.529)	(0.558)	(0.762)	
INTENSITY(LAG)	4.319***	4.040***	4.280***	4.324***	4.192***	3.892***	0.694***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
R^2	0.686	0.663	0.697	0.686	0.693	0.683	
Obs	4105	3688	4030	4036	4012	3448	4361

Note. The dependent variable is INTENSITY. Models 1–6 contain country and year dummies while column 7 contains country-year dummies. Estimation is by maximum likelihood in an ordered logit model except column 7 which is estimated by OLS. p-values based on robust standard errors clustered at the country level are in parentheses. $^{*}p < 10$, $^{**}p < .05$, $^{***}p < .01$.

A.2. **Analysis of correlation vs causation.** This section takes up the issues of correlation vs causation that are discussed in Section 5 of the paper. The analysis has three parts. Section A.2.1 considers the relationship between WGI and other group-level variables. Section A.2.2 presents a test for omitted variable bias. And section A.2.3 considers the issue of reverse causation.

A.2.1. *Group-level variables relationship with WGI.* As discussed in Section 5.1 in the main text, this section treats within-group inequality as the dependent variable and explores the correlation of this variable with other group-level variables. We begin by considering group-level variables already included in the baseline regression models (see Table 2). Model 1 in Table A.6 presents the results where a country-group is the unit of analysis, G^R is the *dependent* variable, and the right-hand side variables include the group-level variables in column 5 of Table 2, as well as country fixed effects. The only variable with a precisely estimated coefficient is group GDP: not surprisingly, groups that are on average richer tend to be *less* unequal. It is worth noting that the analysis reveals no relationship between horizontal inequality and within-group inequality.

Next we examine whether other variables that have *not* been the focus of previous conflict studies have a systematic relationship with G^R. The first type of variable is cultural heterogeneity within a group: ethnic or religious divisions might be related to inequality within a group (for instance, if one subgroup asserts economic advantages over another) and might also be related to a group's cohesion and its ability to coordinate fighting. Model 2 includes LINGUISTIC FRAC., a variable measuring the linguistic fractionalization within a group, and model 3 includes RELIG. FRAC., a variable measuring the religious fractionalization within a group.[36] The results suggest no relationship between the group Gini and religious divisions within a group, but the coefficient for LINGUISTIC FRAC. is positive and highly significant, raising the possibility that if linguistic fractionalization has a negative effect on civil conflict intensity, our estimates of the group Gini coefficient in the conflict regressions might be biased downward. We address this issue in Table A.7 below.

[36] LINGUISTIC FRAC. is based on the group language variables in the EPR data set and is set equal to $1 - \text{size}^2_{LG_1} - \text{size}^2_{LG_2}$, where size_{LG_1} and size_{LG_2} indicate the size of the two largest linguistic subgroups within the group. Religious fractionalization is defined similarly.

Table A.6. Group level variables related to WGI

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
LINGUISTIC FRAC.		0.054**							0.052**
		(0.026)							(0.026)
RELIG. FRAC.			-0.032						-0.049
			(0.369)						(0.136)
REGIONAL				0.003					0.004
				(0.913)					(0.885)
URBANIZED					-0.003				0.002
					(0.825)				(0.886)
RAINFALL						0.033			0.036
						(0.220)			(0.209)
AGRICULTURE							0.000		-0.000
							(0.868)		(0.628)
DIST. CAP								-0.000	-0.000
								(0.571)	(0.371)
HI(LN)	0.002	0.002	0.002	0.002	0.002	0.003	0.002	0.002	0.002
	(0.590)	(0.575)	(0.740)	(0.587)	(0.620)	(0.486)	(0.584)	(0.601)	(0.655)
GROUP GDP(PWT)	-0.216***	-0.206***	-0.213***	-0.216***	-0.227***	-0.224***	-0.228***	-0.229***	-0.210***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
GROUP SIZE	0.041	0.045	0.039	0.042	0.041	0.029	0.039	0.037	0.033
	(0.141)	(0.101)	(0.147)	(0.186)	(0.218)	(0.322)	(0.209)	(0.232)	(0.258)
GROUP ELEV. (SD)	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.521)	(0.496)	(0.520)	(0.527)	(0.526)	(0.615)	(0.591)	(0.571)	(0.638)
GROUP DIAMONS	-0.005	-0.011	-0.004	-0.005	-0.003	-0.005	-0.004	-0.004	-0.009
	(0.732)	(0.480)	(0.765)	(0.730)	(0.832)	(0.785)	(0.833)	(0.793)	(0.617)
GROUP OIL	0.007	0.002	0.010	0.007	0.012	0.016	0.012	0.013	0.017
	(0.681)	(0.896)	(0.541)	(0.685)	(0.535)	(0.439)	(0.546)	(0.502)	(0.325)
GIP	-0.014	-0.014	-0.015	-0.015	-0.010	-0.008	-0.010	-0.009	-0.008
	(0.348)	(0.333)	(0.349)	(0.332)	(0.527)	(0.630)	(0.545)	(0.546)	(0.607)
c	1.698***	2.349***	2.437***	1.694***	2.732***	2.465***	2.737***	2.783***	2.395***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
R^2	0.782	0.791	0.782	0.782	0.783	0.792	0.784	0.784	0.805
Obs	245	244	244	245	235	235	235	235	234

Note. The dependent variable is G^R. All models contain country fixed effects. Estimation is carried out by OLS. P-values based on robust standard errors clustered at the country level are in parentheses. $^{*}p < 10$, $^{**}p < .05$, $^{***}p < .01$.

Two others variables related to the geographic dispersion of groups might at once be related to group inequality and conflict. A group that is regionally concentrated may have lower levels of inequality than groups that are spread out if this regional concentration constrains the nature of economic opportunities. It is also plausible that regionally concentrated groups may find it easier to coordinate fighting, and thus to sustain conflict, creating concerns about omitted variable bias. Model 4 therefore includes REGIONAL, an indicator variable that takes the value 1 if a group is regionally concentrated.[37] This variable has no precise relationship with a group's Gini. Similarly, a more urbanized group may have higher levels of inequality than a group that is more rural (if the nature of economic opportunities varies more in cities) and may have different propensities to sustain conflict (if for example, it is more difficult to sustain conflict in urban areas). Model 5 includes URBANIZATION, a variable that measures the proportion of a group's cells that are urbanized.[38] Again we find no relationship between this variable and a group's Gini.

Finally, we consider two variables related to the terrain a group occupies. First, groups in areas with more rainfall could have more inequality if adequate rainfall creates more opportunities for more productive economic activity among skilled and industrious individuals, and groups in such areas might also be associated with more conflict for reasons unrelated to labor-captial considerations (if such areas are simply attractive to plunder by the government or by other groups). RAINFALL measures the average annual rainfall in the areas controlled by a group, and it is included in model 6. We find no relationship between this variable and WGI. For reasons related to those regarding precipitation, groups that live in areas that can be used for agriculture might have more group-based inequality and might be attractive to governments or other groups. Model 7 therefore includes AGRICULTURE, which measures the percentage of a group's area that is covered with agricultural production. Again, we find no relationship between this variable and the group Gini.[39] Finally, model 8 presents the results when all of these additional variables are included, and the findings reinforce those above: only group GDP and group language fractionalization have a robust association with a group's Gini.

Since our regressions with standard controls include the group's GDP, the analysis in Table A.6 identifies one additional variable that raises concern about omitted variable bias: linguistic fractionalization. This variable is not included in the standard set of group-level controls, it has a precisely estimated relationship with the group Gini, and it could plausibly influence the ability

[37] This variable is taken from the EPR data set and equals 1 if geo_typename equals "Regionally based."

[38] All the variables that are introduced in the remainder of this section are taken from the recently released PRIO-GRID 2.0 (Tollefsen, Bahgat, Nordkvelle and Buhaug 2016. URBANIZATION is the variable in this dataset called urban_gc_mean.

[39] RAINFALL is the log of $prec\_gpcp\_mean$ in Tollefsen et al and AGRICULTURE is $agri\_gc\_mean * 100$ in this data set.

Table A.7. Group linguistic fractionalization and civil conflict

	(1)	(2)	(3)	(4)	(5)
G^R	5.826**	6.114*	7.137**	7.668**	5.629**
	(0.013)	(0.052)	(0.034)	(0.039)	(0.014)
HI(LN)	0.004	-0.001	0.026	-0.007	0.017
	(0.983)	(0.995)	(0.892)	(0.973)	(0.942)
LINGUISTIC FRAC.	0.614	0.599	0.573	0.281	0.395
	(0.582)	(0.600)	(0.602)	(0.795)	(0.715)
GROUP SIZE	-1.436	-1.531	-1.744	-1.001	-0.496
	(0.495)	(0.461)	(0.449)	(0.702)	(0.834)
GROUP ELEV. (SD)	-0.002	-0.002	-0.002	-0.001	-0.001
	(0.552)	(0.514)	(0.460)	(0.584)	(0.711)
GROUP DIAMONDS	0.999	0.994	1.112	0.893	0.883
	(0.389)	(0.386)	(0.333)	(0.437)	(0.451)
GROUP OIL	0.503	0.499	0.548	0.351	0.356
	(0.357)	(0.358)	(0.327)	(0.549)	(0.545)
GROUP GDP		0.122	0.607	0.849	
		(0.861)	(0.446)	(0.332)	
POP			2.875	-1.682	-1.643
			(0.461)	(0.484)	(0.498)
GDP			-1.189	0.131	0.917
			(0.347)	(0.941)	(0.531)
XPOLITY				-0.124*	-0.124*
				(0.063)	(0.063)
EXCLUDED				0.868	0.810
				(0.195)	(0.255)
INTENSITY(LAG)	4.095***	4.098***	4.065***	4.286***	4.301***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
R^2	0.661	0.661	0.661	0.688	0.688
Obs	4448	4448	4358	4154	4154

Note. The dependent variable is INTENSITY. All models contain country and year dummies. This table replicates columns 3–7 in Table 2 introducing LINGUISTIC FRAC, the only variable that turned out to be significantly associated to WGI in Table A.6 above. Estimation is by maximum likelihood in a conditional logit model. p-values based on robust standard errors clustered at the country level are in parentheses. $^{*}p < 10$, $^{**}p < .05$, $^{***}p < .01$.

of a group to sustain fighting for reasons unrelated to the ER theoretical model. We therefore added linguistic fractionalization to models 4-7 in Table 2, and the results Table A.7 show no association between linguistic fractionalization and conflict. The coefficient for the group Gini remains positive and precisely estimated across all five models. Thus, there is little evidence that the results in the preceding discussion are not robust to the inclusion of this variable.[40]

[40]We also added each of the other variables discussed above – RELIG. FRAC., REGIONAL, URBANIZATION, RAINFALL and AGRICULTURE – to models 3-7 – and the results for the group Gini remain highly robust.

A.2.2. *Employing Oster (2016) to assess the possible influence of omitted variables.* To test for the possible importance of omitted variable bias, we have computed the amount of correlation between the unobservables and WGI, relative to the correlation of the observables and WGI, that would be necessary to explain away our key result (i.e., to make the coefficient of WGI equal to zero). In its simplest formulation, this value, denoted by δ, can be computed as follows (see Oster, 2016):[41]

$$\delta = \frac{\beta_c}{\beta_{nc} - \beta_c} \frac{R_c^2 - R_{nc}^2}{R_{\max}^2 - R_c^2},$$

where β_c and β_{nc} are the coefficients of WGI in a model that contains all the observable controls and one with no or a few controls, respectively, and R_c^2 and R_{nc}^2 are the R^2's associated with those regressions. Finally, $R_{\max}^2$ is one's assumption about the maximum R^2 that could be attained if all the relevant controls were observed.

Table A.8. Assessing the importance of omitted variable bias

	RESTRICTED MODEL I	RESTRICTED MODEL II	RESTRICTED MODEL III
$R_{max} = .8$	202.8	189.6	1.65
$R_{max} = .85$	150.7	140.9	1.23
$R_{max} = 9$	119.9	112.1	0.98

Note. This table applies Oster (2016) technique to assess how strong the correlation between the unobservables and WGI, relative to the correlation of the observables and WGI, has to be in order to explain away the significance of WGI. Calculations have been performed using the software *psacalc* provided by the author. The full model corresponds to column 7 in Table 2, and the restricted models contain no controls (I), year dummies (II) and country and year dummies (III). We consider three different values of $R_{max} = \{0.8, 0.85, 0.9\}$, the maximum R^2 that could be attained in all the relevant controls would be introduced in the regression. See the main text and Oster (2016) for details.

[41]This definition of δ corresponds to the case where there is a single observable control, see Oster (2016) for details on the more general case

A value of $\delta = 2$, for example, would suggest that the unobservables would need to be twice as important as the observables to produce a treatment effect of zero. Altonji et al. (2005) and Oster (2016) suggest that values of δ larger than 1 in absolute value can be interpreted as evidence that omitted variable bias is unlikely to explain the observed result. A value of 1 (or larger) means that the unobservables would need to be at least as important as the observables to produce a treatment effect of zero. Since researchers typically choose the controls they believe *ex ante* to be the most important (Angrist and Pischke, 2010), situations where the effect of the unobservables is larger than that of the controls are deemed unlikely.

Table A.8 presents the results. The full model corresponds to that in column 7 in Table 2. Restricted models I, II and III correspond to models with no controls, with year fixed effects and with country and year fixed effects, respectively. In order to implement the test, a value for R^2_{max} needs to be chosen. To select this value, we have followed the advice in Oster (2016) who suggests using a value equal to 1.3 times the value of the R^2 obtained in the regression with all controls, which in our case equals 0.85. In addition, we have also considered two additional values, so that $R^2_{\text{max}} = \{0.8, 0.85, 0.9\}$. The figures in Table A.8 correspond to the values of δ for each of the 9 cases considered. In most cases we obtain values of δ that are larger than 1 in absolute value, which suggests that it is not likely that the significance of WGI is due to omitted variable bias. Only in one of the nine combinations (when we consider a model with country and year fixed effects – Restricted model I – and a very high value of R_{max} (equal to 0.9) do we obtain values of δ that are smaller than 1. This test therefore suggests that it is very unlikely that the results for within-group inequality are driven by omitted variables.

A.2.3. *Reverse causation.* The test for reverse causation is described in Section 5.3. Table A.9 presents the results.

Table A.9. Reverse causation: Civil conflict and changes in WGI

%WAR YRS	-0.014	0.029	0.022	0.079	0.009	0.045	0.157	0.083
	(0.814)	(0.687)	(0.932)	(0.786)	(0.785)	(0.511)	(0.555)	(0.744)
TOTAL YEARS	-0.001	-0.003	-0.000	0.001***	-0.009**	-0.020***	-0.009**	-0.016***
	(0.722)	(0.293)	(0.935)	(0.000)	(0.016)	(0.000)	(0.028)	(0.000)
LAST YEAR	-0.001	-0.011***	-0.001	-0.010***	0.014**	0.020***	0.014**	0.021***
	(0.858)	(0.000)	(0.810)	(0.000)	(0.015)	(0.000)	(0.018)	(0.000)
CONSTANT	1.856	21.726***	2.561	19.747***	-27.774**	-40.302***	-27.987**	-42.127***
	(0.858)	(0.000)	(0.811)	(0.000)	(0.015)	(0.000)	(0.019)	(0.000)
Obs	251	251	243	243	251	251	243	243

Note. The unit of observation is a group and the dependent variable in models 1-4 (5-8) is the change in G^R (G^I) between the first year and last year for which we have different surveys measuring group inequality. %WAR YRS is the proportion of years that the group is in conflict over the period for which we have data for the group. TOTAL YEARS is the number of year from the first year for which we have group data to the last year for which we have group data. LAST YEAR is the last year for which we have group data. Models 1,2, 5 and 6 use all data whereas models 3, 4, 7 and 8 exclude groups that were in conflict during the first year for which we have surveys. Models 3, 4, 7 and 8 include country fixed effects whereas the other models do not. Models are estimated using OLS, and p-values based on robust standard errors clustered at the country level are in parentheses. $^{*}p < 10$, $^{**}p < .05$, $^{***}p < .01$.

A.3. **WGI and conflict onset: Additional analysis.** This section revisits the analysis in Kuhn and Weidmann (2015), as discussed in the paper's Section 6.1. Model 1 in Table A.10 replicates model 4 in Table 1 of Kuhn and Weidmann's paper, which relies on the same set of control variables as in CWG, but which (unlike CWG) includes country fixed effects. Since we are considering all years (starting in 1992) for which EPR data is available, our data merging process results in more observations than in Kuhn and Weidmann's original paper.[42] However, the results are similar to those presented in Kuhn and Weidmann: the coefficient for IGI is positive and significant at the 10% level, though not as precisely estimated as in their paper.[43]

This result for IGI, however, is not at all robust. Omitted variable bias is one concern. We explored the robustness of the results to omitted variables employing the technique developed by Oster (2016), as in Table A.8 in a setup where the full model is that in Kuhn and Weidmann's

[42]The data employed for this regression has been taken from the Growup portal (`https://growup.ethz.ch`) except for IGI, which has been computed by Kuhn and Weidmann (2015).

[43]The models presented in the previous tables use lags of the economic variables in order to diminish concerns about reverse causality. For consistency with the Kuhn and Weidmann (2015) and CWG approaches, variables in Table A.10 are not lagged.The results are substantively similar regardless of whether the variables are lagged. Model 1 also omits 15 observations by using exclusion rules commonly used in research that employs EPR data to estimate models of civil war onset. In particular, groups are excluded if they are judged to be dominant, to have a monopoly on power, or if they are geographically dispersed (see discussion in CWG 2011). The results are essentially identical if these exclusion rules are ignored.

(2015) (column 1 in Table A.10) and the three restricted models contain (i) no controls, (ii) country fixed effects and (iii) country and year dummies.[44] In all cases the values of δ are close to zero, raising concerns that the estimates for IGI in model 1 could suffer omitted variable bias.[45]

This concern is clear from the model in column 2, which adds the group-level control variables that are present in the models in Section 4 and that are taken directly from the *Growup* portal: GROUP GDP, GROUP ELEV.(SD), GROUP SIZE, GROUP DIAMONDS and GROUP OIL.[46] Though the coefficient for IGI remains positive, it is now estimated with considerable error (p=.30). Of the five group-level variables added to model 2, however, only one is significant, GROUP ELEV. (SD), so model 3 re-estimates model 2 omitting the insignificant group level regressors. Model 1 also lacks two country-level controls, POP and POLITY. Column 4 adds these two variables and shows POP has a positive and significant coefficient while POLITY is estimated with a large error. The coefficient for IGI decreases to less than half the value obtained in Column 1 and remains very imprecisely estimated (p-value .42). Finally, it is useful to note that the results in model 1 require the presence of a particular group: the East Timorese in Indonesia. Model 5 presents results from re-estimating model 1 without this group and the coefficient for IGI is insignificant. Columns 6 and 7 reproduce columns 3 and 4 omitting this group, obtaining similar results.

We therefore find little support for a robust association between within-group inequality and conflict onset, regardless of whether WGI is measured using nightlights data or surveys. While it is always possible that these null results are due to measurement error, we suspect that is not the issue here. Instead, the null results are not inconsistent with the ER argument, which emphasizes that WGI increases the capacity to fight rather creating incentives to do so.

---

[44]Kuhn and Weidmann (2015) also consider in their robustness checks an omitted variable bias analysis in a similar vein as the one discussed in the text, and obtain values of δ that suggest that results are robust to omitted variable bias. However, they use the technique introduced by Bellows and Miguel (2009) that does not incorporate the movements in R^2 in the estimation of δ. As discussed at length in Oster (2016), the omitted variable bias is proportional to the movement in coefficients *only* if movements in R^2's are also taken into account and, thus, it is critical to introduce this term in order to have accurate results.

[45]As in Table A.8 in Appendix A, we have considered different values for R^2_{max}, the maximum value of the R^2 coefficient that could be obtained if all the relevant controls were included in the regressions. Only when R^2_{max} was set as low as 0.1 did we obtain values δ larger than 1 in some specifications.

[46]GROUP SIZE is SIZE (EPR) from the EPR data, and is the "group's population size as a fraction of the ethnically relevant population of this group's country."

Table A.10. WITHIN-GROUP INEQUALITY (MEASURED USING NIGHTLIGHTS) AND CONFLICT ONSET

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
IGI	1.921*	1.171	1.456	0.900	1.581	1.089	0.752
	(0.082)	(0.301)	(0.222)	(0.422)	(0.181)	(0.393)	(0.513)
LINEQ2	0.210	0.356	0.216	0.236	-0.649	-0.646	-0.511
	(0.208)	(0.163)	(0.258)	(0.179)	(0.296)	(0.332)	(0.493)
EXCLUDED	1.002**	0.993**	1.046**	1.224**	0.993**	1.051**	1.096**
	(0.017)	(0.019)	(0.011)	(0.012)	(0.023)	(0.014)	(0.034)
POW. BALANCE	-2.465	-2.854	-3.938	-2.906	-2.705	-4.327	-4.225
	(0.435)	(0.467)	(0.286)	(0.434)	(0.384)	(0.237)	(0.333)
POW. BALANCE^2	1.509	2.594	3.212	4.422	1.569	3.437	3.607
	(0.695)	(0.658)	(0.465)	(0.488)	(0.688)	(0.437)	(0.503)
GDP	-0.105	-0.278	-0.099	-0.543	-0.198	-0.204	-0.555
	(0.902)	(0.701)	(0.909)	(0.681)	(0.817)	(0.814)	(0.677)
N. EXCL. GROUPS	-0.001	0.004	0.006	0.051	0.025	0.030	0.202
	(0.988)	(0.944)	(0.919)	(0.693)	(0.741)	(0.697)	(0.319)
GROUP GDP		0.321					
		(0.395)					
GROUP ELEV. (SD)		0.001*	0.001**	0.001		0.001**	0.001
		(0.089)	(0.042)	(0.133)		(0.041)	(0.134)
GROUPSIZE		-0.463		-2.153			
		(0.898)		(0.569)			
GROUP OIL		0.243					
		(0.643)					
GROUP DIAMONDS		-0.589					
		(0.181)					
POP				0.805**			0.756*
				(0.044)			(0.069)
XPOLITY				-0.029			-0.011
				(0.781)			(0.926)
YEAR	-0.071	-0.072*	-0.076	-0.045	-0.067	-0.072	-0.045
	(0.126)	(0.097)	(0.103)	(0.450)	(0.169)	(0.140)	(0.474)
PEACEYRS	-0.284	-0.251	-0.280	-0.256	-0.300*	-0.291	-0.276
	(0.106)	(0.175)	(0.114)	(0.225)	(0.085)	(0.100)	(0.183)
SPLINE1	0.010	0.008	0.010	0.007	0.012	0.011	0.009
	(0.483)	(0.614)	(0.511)	(0.669)	(0.422)	(0.467)	(0.582)
SPLINE2	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.691)	(0.830)	(0.729)	(0.887)	(0.616)	(0.670)	(0.790)
SPLINE3	0.000	0.000	0.000	-0.000	0.000	0.000	0.000
	(0.817)	(0.952)	(0.854)	(0.974)	(0.734)	(0.787)	(0.927)
R^2	0.192	0.202	0.197	0.224	0.183	0.189	0.213
Obs	2982	2977	2977	2591	2977	2972	2586

Note. The dependent variable is ONSET. All models contain country dummies. A conditional logit estimator has been employed. p-values based on robust standard errors clustered at the country level are in parentheses. $^{*}p < 10$, $^{**}p < .05$, $^{***}p < .01$.

APPENDIX B. GROUP INEQUALITY DATA

This Appendix provides additional details about the survey-based dataset, as well as other data used in the analysis. Section B.1 describes the set of surveys used, how these surveys are employed to measure a respondent's ethnic identity and income, and how the income and identity measures are used to measure the inequality variables. Section B.2 then describes how we adjust the group Ginis using the ratio and intercept approach. Section B.3 discusses existing measures of horizontal inequality and their limitations, while Section B.4 presents summary statistics of the survey-based ethnic inequality data. Section B.5 discusses the strengths and weaknesses of using survey as opposed to geo-coded data for measuring group inequality. In so doing, we assess issues regarding the representatives of the countries in the survey data set, the representativeness of groups within surveys, and the reliability of the survey-based income measures. We also argue that the geo-coded data is much more useful for measuring inequality across groups than inequality within groups. Finally, section B.6 describes how we use the EPR data to create measures of horizontal inequality.

B.1. **The survey data.** As noted in the main text, the surveys used include "HES" (for "Household Expenditure Survey") surveys, the World Values Surveys (WVS), and the Comparative Study of Elections Surveys (CSES) (which reports income in quintiles), which report some form of household income. We also use the Demographic Health Surveys (DHS) and the Afrobarometer, which are typically conducted in relatively poor countries and do not have direct measures of household income data, but rather have information on various assets that households possess. In these countries, where many poor individuals do not make substantial cash transactions, social scientists often use asset indicators to determine household "income." For these surveys, we detail below how the household assets are used to this end.

We have been able to compute Gini coefficients for 446 groups in 89 countries. Figure B.2 depicts the countries for which we have data. Figure B.2 shows the geographic distribution of countries in the data set and Table B.11 lists the surveys used to create the data set.

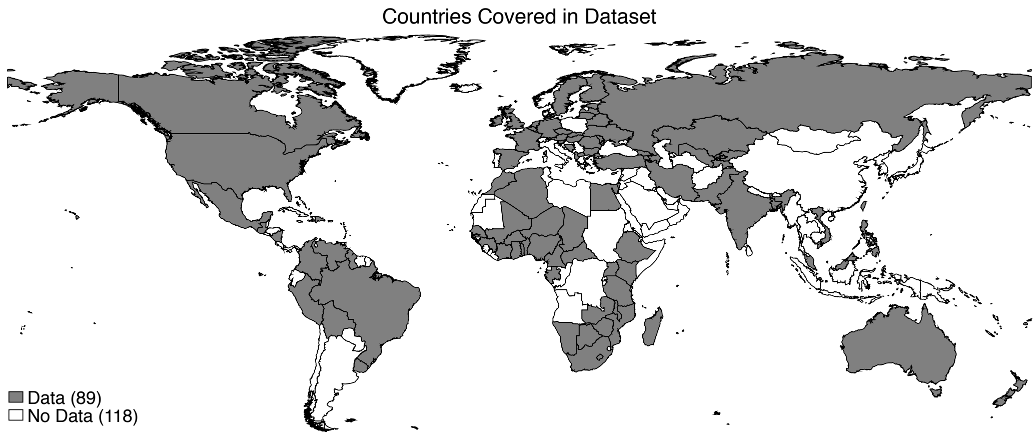


Figure B.2. Countries included in data set

Group identity of respondents. Our definition of groups combines information from Fearon (2003) and Cederman, Wimmer and Min (2010). The Fearon definition of groups emphasizes seven dimensions of group identity that include cultural attributes, an identity inherited at birth, and a collective history. For our purposes, a key feature of using Fearon's definition is that the definition is not a function of political relevance, which can be endogenous to the political dynamics that lead to civil war. The Fearon definition also does not preclude the inclusion of groups based on their geographic characteristics. Thus, Fearon provides important and widely used group-based data base. Using this definition when using the surveys therefore allows us to construct a data set that will be useful not only in the present study of conflict, but also in future studies that use the data for other reasons.

The surveys typically include one or more of four types of questions that are relevant (depending on the country and on Fearon's definition) to coding an individual's ethnic identity: ethnicity, language, race and religion. Mapping from the categories included in the surveys to the categories defined by Fearon is often straightforward. In Macedonia, for example, Fearon lists five groups: Macedonian, Albanian, Turk, Roma and Serbs, and in WVS, each of these groups is a unique category in the ethnicity variable (x051a). In Bangladesh, Fearon lists two groups, Bengali and Hindi and we have a DHS survey from 1997 in Bangladesh. The DHS survey has a religion variable where 89.7 percent of respondents are Muslim, 0.26 percent are Buddhist, 0.16 percent

Table B.11. The Surveys

Albania	2002(WVS) 2005(HES-LSMS)	Kyrgyz Rep	1997(DHS) 2003(WVS)
Algeria	2002(WVS)	Latvia	1996(WVS) 1999(WVS)
Armenia	1997(WVS) 2000(DHS)	Lithuania	1997(CSES, WVS)
Australia	1995(WVS) 1996(CSES) 2004(CSES) 2005(WVS)	Macedonia	1998(WVS) 2001(WVS)
Austria	2000(LIS)	Madagascar	2005(AFRO)
Azerbaijan	1995(HES-ASLC) 1997(WVS) 2006(DHS)	Malawi	2000(DHS) 2003(AFRO) 2004(DHS) 2005(AFRO)
Bangladesh	1996(WVS) 1997(DHS) 2000(DHS) 2002(WVS) 2004(DHS) 2007(DHS)	Malaysia	2006(WVS)
Belarus	1996(WVS) 2001(CSES)	Mali	1995(DHS) 2001(DHS) 2002(AFRO) 2005(AFRO) 2006(DHS)
Belgium	1999(CSES, WVS)	Mexico	1997(CSES, WVS) 2000(WVS) 2003(CSES)
Benin	1996(DHS) 2001(DHS) 2005(AFRO) 2006(DHS)	Moldova	1996(WVS) 1999(WVS) 2005(DHS) 2006(WVS)
Bolivia	2002(HES-MECOVI) 2003(DHS)	Morocco	2001(WVS) 2007(WVS)
Bosnia	1998(WVS) 2001(WVS) 2004(HES-LIBP)	Mozambique	2002(AFRO) 2005(AFRO)
Botswana	2003(AFRO) 2005(AFRO)	Namibia	2000(DHS) 2003(AFRO) 2006(AFRO)
Brazil	1996(DHS) 1997(WVS) 2002(CSES, HES-IPUMS) 2006(WVS, HES-PNAD)	Netherlands	1999(WVS)
Bulgaria	1995(HES-IHS) 1997(WVS) 2001(CSES) 2006(WVS)	New Zealand	1996(CSES) 1998(WVS) 2002(CSES)
Burkina Faso	1992(DHS) 1998(DHS, HES-EP2) 2003(DHS)	Nicaragua	2001(HES-EMNV)
Cameroon	1998(DHS) 2004(DHS)	Niger	1992(DHS) 1998(DHS) 2006(DHS)
Canada	1997(CSES, HES) 2000(WVS) 2001(HES-IPUMS) 2006(WVS)	Nigeria	2000(WVS) 2005(AFRO)
Central African Rep	1994(DHS)	Pakistan	2001(WVS)
Chad	1997(DHS) 2004(DHS)	Peru	2000(DHS) 2004(DHS, HES) 2008(WVS)
Colombia	1998(WVS)	Philippines	1993(DHS) 1998(DHS) 2003(DHS) 2008(DHS)
Cote d'Ivoire	1998(DHS)	Romania	1996(WVS, CSES) 1997(HES) 2005(WVS)
Cyprus	2006(WVS)	Russia	1995(WVS) 1999(CSES) 2000(CSES, HES) 2006(WVS)
Czech Rep	1996(CSES)	Senegal	1992(DHS) 2002(AFRO) 2005(AFRO, DHS)
Dominican Rep	1998(WVS)	Singapore	2002(WVS)
DRC	2007(DHS)	Slovakia	1998(WVS)
Egypt	1995(DHS) 2000(WVS) 2005(DHS) 2008(DHS)	Slovenia	1996(CSES)
Estonia	1996(WVS) 1999(WVS) 2000(HES)	Spain	1995(WVS) 1996(CSES) 2000(CSES, WVS) 2004(CSES) 2007(WVS)
Ethiopia	2000(DHS) 2005(DHS)	South Africa	1996(WVS) 1998(DHS) 2001(HES-IPUMS) 2002(AFRO) 2006(AFRO) 2007(WVS)
Finland	2003(CSES) 2004(HES) 2005(WVS)	Sweden	2005(HES) 2006(WVS)
France	1999(WVS) 2002(CSES) 2006(WVS)	Taiwan	1995(WVS) 1996(CSES) 2004(CSES)
Gabon	2000(DHS)	Tajikistan	1996(HES-LSS)
Georgia	1996(WVS)	Tanzania	1993(HES-HRDS)
Germany	1999(WVS) 2004(HES) 2006(WVS)	Togo	1998(DHS)
Ghana	1993(DHS) 1998(DHS) 2003(DHS) 2008(DHS)	Turkey	1993(DHS) 2007(WVS)
Guatemala	1995(DHS) 1998(DHS) 2000(HES-ENCOVI) 2005(WVS) 2006(HES)	Uganda	1995(DHS) 2005(AFRO)
Guinea	1999(DHS) 2005(DHS)	UK	2004(HES)
Guyana	2005(DHS)	Ukraine	1996(WVS) 1998(CSES) 2006(WVS)
Hungary	2002(CSES)	United States	1996(CSES) 1997(HES) 2000(WVS) 2004(CSES) 2005(HES-IPUMS) 2006(WVS)
India	1995(WVS) 2001(WVS) 2006(WVS)	Uruguay	1996(WVS) 2006(WVS)
Iran	2007(WVS)	Uzbekistan	1996(DHS)
Ireland	1999(WVS)	Venezuela	1996(WVS) 2000(WVS)
Israel	1995(HES-IPUMS) 2005(HES)	Vietnam	1997(DHS) 2002(DHS) 2005(DHS)
Kazakhstan	1995(DHS) 1999(DHS)	Zambia	1996(DHS) 2001(DHS) 2003(AFRO) 2005(AFRO) 2007(WVS, DHS)
Kenya	1993(DHS) 1998(DHS) 2003(DHS, AFRO) 2005(AFRO) 2008(DHS)	Zimbabwe	2001(WVS) 2004(AFRO) 2005(AFRO)

Note. This table presents all the surveys available and their corresponding dates for each of the countries in our data.

are Christian and 9.91 percent are Hindu. We use this variable to code the Hindus, and the Bengalis are coded as the Muslims. As a practical matter, the coding of the Buddhists and Christians is irrelevant because they are a trivial percentage of the population. We discard surveys that do not allow us to adequately identify the Fearon groups. Specifically, if there exist one or more groups on Fearon's list that we cannot identify in the survey, we sum the proportion of the population that these groups represent per Fearon's data. If this sum is greater than .10, we do not utilize the survey. As an example, consider the Afrobarometer survey for Nigeria in 2003, for which it is possible to use a language variable to map to many of Fearon's groups. But one of Fearon's groups is "Middle Belt," and it is not possible to identify these individuals in the Afrobarometer survey. Since Fearon's data suggest they represent 18 percent of the population (which exceeds our threshold), we exclude this survey. The DHS surveys include multiple members from the same household, which can present special issues related to categorizing the ethnic identity of mixed households. For such households, we use the ethnicity that is most common, and if there is no modal ethnicity, we use the ethnicity of the head of household.

Data on the participation of groups in conflict comes from the definition of groups used in the Ethnic Power Relations data. A discussion of the definition of groups is found in Vogt, et al. (2015) and Cederman, Wimmer and Min (2010). For the purposes of building a general data set on group inequality, the definition of groups is less attractive than Fearon because the EPR definition includes "political relevance." That is, for a group to exist, it must have a political organization that advocates politically for the ethnic group, or it must be subject to political discrimination. This definition can therefore introduce an unwanted link between the definition of a group and the dynamics of the politics being studied, making difficult, for instance to understand why some groups become politically active while others do not (see Huber 2017). But the definition has the advantage of being linked to data on group participation in conflict, as described in Wucherpfennig, Metternich, Cederman and Gleditsch (2012). And as a practical matter, the two definitions of groups are very closed related and can be combined.[47] Thus, to combine our data on group economic attributes with data on conflict, we map the Fearon definitions to the EPR definitions. Details of this mapping are in the replication files.

Income. For the data sets with direct measures of income, we use the income reported by the respondent. As noted in the main text, the income metrics vary across survey types, and are most detailed for the HES surveys. For the data sets without incomes, we use responses to questions about living conditions to construct measures of household "income." The Afrobarometer surveys

[47] See, for example, Cederman, Weidmann and Bormann (2015).

(which represent less than 10 percent of all surveys used) have a relatively small number of asset questions (typically 5 or less), all of which are indicator variables. For these surveys we simply sum the assets. For the DHS surveys, which contain a large number of asset indicators (typically around 13), we follow Filmer and Pritchett (2001) and McKenzie (2005) and run a factor analysis on the asset variables to determine the weights of the various assets in distinguishing household well-being. We then use the factor scores, and the responses to the asset questions, to measure the household "wealth" of the respondent.

More specifically, the DHS input variables included several categorical variables: source of household drinking water, type of toilet facility, type of floor material, and type of roof material.[48] The categories for these variables are recoded so that higher number reflects more desirable conditions. For example, for type of toilet facility, a covered latrine is coded as 4, an open latrine as 3, a septic pit as 2 and no facility as 1. Uninterpretable categories (like "other") are recoded as missing. The DHS surveys also often include a variable tapping the number of sleeping rooms. And there are a number of indicators marking access to certain types of "equipment": electricity, a radio, television, a refrigerator, a bicycle, a scooter, a car and a phone. All of these variables were used to conduct a principal components analysis, which produce weights on the input variables (the factor scores). These weights and the survey responses are used to construct an "income" for each household. Since there is variation across surveys regarding which input variables are included, to make the incomes comparable across DHS surveys, we convert the "incomes" to percentile ranks, which became the inputs for constructing the inequality variables.

Computing the Horizontal and Within-group Inequality Measures. Once the ethnic identity and "income" of each respondent is determined, we calculate a group's GDP per capita (the main input needed in the elaboration of the HI measures) and the Gini coefficient for each group. To enhance the representativeness of the surveys, when survey weights exist (which is the case for most surveys), we expand the data set using this weight before calculating the measures. To calculate a group's GDP, we use a survey's information about group income and group size to measure a group's total share of income. For example, suppose there are two groups, and group

[48]For DHS, household income was calculated using the "household" data set (rather than the male or female data set).

1 has a per capita income of 2 with a size of .2, and group 2 has an per capita income of 5 with a size of .8, then country's income per capita is 4.4 (2*.2+5*.8) and group 1's (2's) share of total income is .09 (.91). For countries where we have multiple surveys, we use the mean of these measures of group shares. We then multiply these measures of group income shares by country GDP per capita (as measured by the Penn World Tables) and divide by group size, yielding a harmonized measure of group GDP per capita. HI measures based on survey data are computed by applying the formulas corresponding to each of the horizontal inequality measures to group's GDP per capita, see below.

To calculate the group's Gini, we use Stata's Ginidesc ado file developed by Aliaga and Montoya (1999). We then adjust these group Ginis, as described in the next section.

B.2. **Adjusting the group Gini coefficients: the intercept and ratio approach.**

The section describes the two approaches used to adjust the group-level Ginis to account for heterogeneity across surveys in the measure of income.

i) The "ratio approach" to adjusting the components of the Gini. The first approach to adjusting the group Gini coefficients draws on external data on the Gini – the Standardized World Income Inequality Dataset (SWIID)– to adjust the group-level measures of the Gini as well as the three components of the Gini decomposition. The SWIID (Solt 2009) provides comparable Gini indices of gross and net income inequality for 173 countries from 1960 to the present and is one of the most thorough attempts to tackle the comparability challenge (see Solt 2009 for details on the methodology). We use the SWIID data and a methodology similar to Solt (2009) to obtain (time-varying) adjustment factors for the overall country Gini from each country and year and then apply these factors to the group Ginis from the surveys. Let $G_{c,t,\text{SWIID}}$ be the SWIID Gini for country c in year t and $G_{c,t,s}$ be the Gini from country c and year t using survey s. The *ratio* approach involves 4 steps. The first three are similar to employed by Solt (2009) to adjust country-level inequality measures. The fourth step applies the same adjustment procedure to the group-level Ginis.

Step 1: Whenever a survey Gini and the SWIID Gini are available for the same country and year, we compute their ratio, $R_{c,t,s} = \frac{G_{c,t,s}}{G^S_{c,t}}$.

Step 2: For the 201 available ratios, we regress $R_{c,t,s}$ on country (α_c) and year (δ_t) dummy variables. Specifically, we estimate:

(1) $$R_{c,t,s} = \alpha_c + \delta_t + \epsilon_{c,t,s}.$$

Step 3. For each survey we use the parameter estimates from eq. (1) to obtain the predicted values of the ratios, $\hat{R}_{c,t,s}$, for all surveys. For those surveys where ratios exist, the predicted ratios are of course very close to the actual ratios (r=.98), but the predicted ratios also can be derived from Eq. (1) for the 16 surveys where the SWIID Gini is missing. This is justified by the fact that the factors that affect these ratios tend to change only slowly over time within a given country and, hence, the missing ratios can be predicted based on available data on the same ratio in the same country in proximate years. As in Solt (2009), the adjusted Gini coefficient for country c, time t and survey s, $\widehat{G_{c,t,s}}$, could be computed as the product of $\widehat{G_{c,t,s}}$= $\overline{R}_{c,t,s}G_{c,t,s}$.

Step 4. To obtain adjusted measures using the ratio approach, we take the product of the group Ginis from the surveys and the predicted ratios:

(2) $$\widehat{G_{g,c,t,s}} = \hat{R}_{c,t,s} * G_{g,c,t,s}.$$

where G_g denotes group-level Ginis. We justify this last step as follows. Remember that the country-level Gini coefficient can be written as:

(3) $$G = WGI_c + BGI_c + OV_c,$$

where WGI_c is the weighted average of group Ginis, BGI_c is a measure of the average difference in group mean incomes in a society and OV_c (from overlap) is a residual term. It follows that

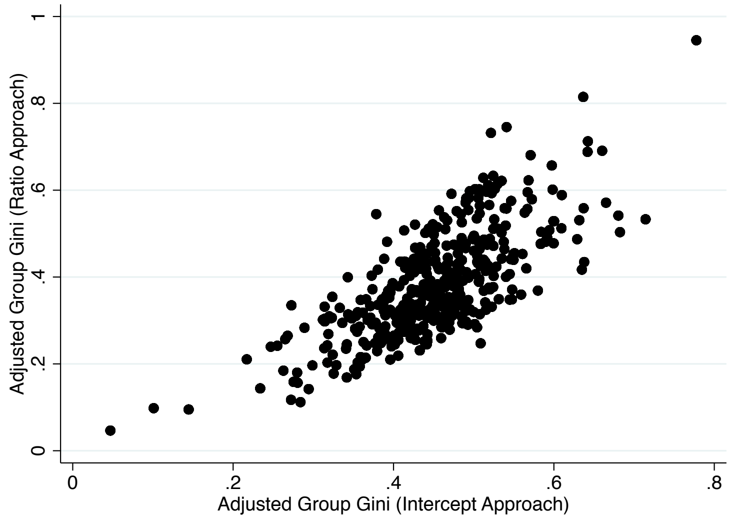


Figure B.3. G^R versus G^I

$\widehat{G_{c,t,s}} = \hat{R}_{c,t,s} * G_{c,t,s} + \hat{R}_{c,t,s} * BGI_{c,t,s} + \hat{R}_{c,t,s} * OV_{c,t,s}$. Since WGI is a weighted average of the group Ginis, it follows that by adjusting each of the group Ginis by the same factor $\hat{R}_{c,t,s}$, one obtains the "adjusted" WGI, $\widehat{WGI} = \hat{R}_{c,t,s} * WGI$. Thus, the adjusted country-level Gini and the adjusted group Ginis are internally consistent.

Step 4 yields the measures we use in our empirical analysis using the "ratio" approach. Such measures are denoted by G^R.

ii) The "intercept approach" to adjusting the survey measures of inequality. The second approach to adjusting the WGI measures is similar to that employed in the original Deininger and Squire (1996) exercise. The idea is to remove average differences due to different survey methodologies. To implement this approach, we first calculate the Gini coefficient for each group using the surveys. We then regress the group Gini on survey, time and country dummies, with HES as the omitted category. We use the HES as reference since these surveys are probably the best-available estimates of income distributions in the world. The shift coefficients on the survey dummies are then used to adjust the inequality measures so as to remove average differences that could be traced to different survey types.

For both approaches, time-invariant measures are computed by averaging all observations available for one group/country and assigning the average values to all years. As noted in the main text, we also compute these measures using only data that exists during or prior to a given year.

Figure B.3 displays G^R vis-à-vis G^I and shows that there is a close relationship between the two sets of survey-based Gini coefficients (the correlation coefficient is .73). G^R has a lower mean (.38) and higher standard deviation (.12) than G^I (.45 and .9, respectively).

B.3. **Existing measures of horizontal inequality and corrected measures.** This section describes measures of horizontal inequality used in previous research and describes the "corrected" measures we employ in the empirical analysis. One such measure is

$$\text{LINEQ2} = (log\frac{G_g}{G})^2,$$

where G_g is the GDP per capita of the group and G is the *unweighted* average GDP per capita of all groups in society.[49] The variable takes a larger value as a group becomes either relatively rich or relatively poor. Thus, if the variable has a robust relationship with civil conflict the relationship cannot be easily attributed to group poverty rather than horizontal inequality.

One flaw with this measure is that G is defined as the unweighted average of all group GDP's per capita. This definition implies that when calculating LINEQ2 for a group j, one is partially comparing that group's wealth to itself (because group j's GDP is included in G). A much more direct measure of grievance would be to compare a group's GDP to that of *other* groups. To see the difference, consider a two-group example. Existing measures will compare the income of group A to the average of the incomes of groups A and B. A more intuitive approach, however, would be to compare the income of group A to that of group B, as the logic of HI suggests.

A second flaw is that the definition of G ignores group size. To see how this can lead to a misleading measure, consider an example with three groups, group A is a poor group with an income of 1, group B is a somewhat less poor group with an income of 1.5, and group C is a rich group with an income of 5. The measure will be the same for group A regardless of whether group

[49]See Cederman et al. (2011) and Kuhn and Weidmann (2015).

C is extremely large (in which case the rest of the population group A faces is relatively poor, like group A) or extremely small (in which case the rest of the population that group A faces is relatively rich, unlike group A).

Finally, it is worth noting that LINEQ2 has a peculiar functional form: the log of a ratio, squared. This highly non-linear function is one way to ensure that the measure grows larger as a group becomes either richer or poorer than average, but it is not a particularly intuitive way to do so: the functional form is extremely sensitive to small changes in the data, and it imposes asymmetries in the treatment of rich and poor groups.[50]

A second approach to measuring horizontal inequality in previous research creates separate variables for rich and poor groups. For example, Cederman, Weidmann and Bormann (2015) define LOW $= max\{1, \frac{G}{G_g}\}$, and HIGH $= max\{1, \frac{G_g}{G}\}$. The goal of developing this measure is to distinguish whether the onset of civil wars is associated more strongly with a group being relatively poor or relatively rich. While these variables avoid the unusual functional form of LINEQ2, they share the other two problems discussed in the previous paragraph.[51] We would also note that when HIGH and LOW are used, arguments about horizontal inequality are supported only if *both variables* are associated with civil war onset. If, for example, only LOW has a significant association, we cannot disentangle the role of poverty from the role of inequality.

To circumvent the above mentioned problems, first we use a "corrected" version of G, $\overline{\mathrm{G}_g}$, which doesn't include group g in its definition and that is sensitive to the relative size of the groups (if there are three or more groups).[52] Second, as discussed in the main text (see Section 3.2) we have introduced new measures that do not impose such large asymmetries in the treatment of poor and rich groups and are less sensitive to small changes in the data.

---

[50]Suppose, for example, there are only two groups in society: a poor group with an income of 1 and a rich group with an income of 5. Then (no matter the size of the two groups) LINEQ2 is .26 for the rich group, and is 1.21 for the poor group.

[51]Fjelde and Østby (2014) invoke a variant of LOW and HIGH that does not include group g in the calculation of G.

[52]As mentioned in Section 3.2 in the main text, $\bar{G}_g$ is defined as total GDP of all groups minus total GDP of G divided by total population in the country minus the population of g. If GDP per capita at the group level is available, $\bar{G}_g$ can also be computed as follows. If there are n groups with sizes $\pi_1 \ldots \pi_n$ and per capita GDPs, $G_i \ldots G_n$, them $\bar{G}_g = \frac{\sum_{i \neq g} \pi_i G_i}{(1-\pi_g)}$.

Our "corrected" versions of LINEQ2 and HIGH/LOW (denoted as $\overline{\text{LINEQ2}}$, $\overline{\text{HIGH}}$ and $\overline{\text{LOW}}$, respectively) are identical to the uncorrected ones but use $\overline{\mathrm{G}_g}$ in place of G.

B.4. **Ethnic Inequality measures: summary statistics.** This section presents a basic description of the ethnic inequality variables based on survey data. Table B.12 presents summary statistics while Table B.13 reports the correlation among the main variables.

Table B.12. Ethnic inequality measures: summary statistics

	Obs	Mean	Std	Skewness	Min	Max
G^R	8474	0.38	0.12	0.64	0.05	0.95
G^I	8474	0.45	0.09	-0.29	0.05	0.78
HI(LN)	8474.00	1.71	3.41	1.52	-4.78	14.43
HI(ABS)	8474.00	11019.97	93537.91	17.32	0.01	1.85e+06
LINEQ2	8474.00	0.10	0.22	6.14	0.00	2.41

Note. This table presents summary statistics of the ethnic inequality variables in our baseline specifications.

Table B.13. Ethnic inequality measures: correlation

	G^R	G^I	LINEQ2	HI(LN)	HI(ABS)
G^R	1				
G^I	0.729	1			
LINEQ2	0.162	0.180	1		
HI(LN)	-0.068	-0.084	0.284	1	
HI(ABS)	-0.052	0.052	0.506	0.383	1

Note. This table presents correlations of the ethnic inequality variables in our baseline specifications.

B.5. **Strengths and weaknesses of survey vs. spatial data.** There are a number of potential limitations associated with using surveys to measure ethnic inequality. One is that the approach can only be implemented in countries with useful surveys, and the set of such countries with such surveys might be unrepresentative in important ways. In particular, one might worry that the countries where surveys exist might be correlated with ethnic conflict itself, or with variables related to ethnic conflict.

Table B.14 examines this issue empirically.[53] The table compares the sample of countries obtained from our surveys to a broader set of countries in the SWIID data set. We use SWIID as a benchmark because it is the inequality dataset with the broadest available coverage. The top half of Table B.14 describes the distribution of countries around the world using the SWIID and our survey data, focusing on the post-1994 time period for which most of our survey data exists. There are 136 countries available in SWIID (taking into account that there are some countries in this data set for which conflict or other control variables do not exist) and 88 countries – or 64 percent of the SWIID – for which we have useful surveys. The table shows a slightly higher proportion of the countries in the survey data are from Central Europe, and a slightly higher proportion of the SWIID countries are from Latin America, but the distributions of countries across the regions are quite similar. Thus, there is little in the way of regional bias in the survey data.

The bottom half of the table provides descriptive data on key variables in the two data sets: GDP/capita, ethnic fractionalization (F), ethnic polarization (P), level of democracy (xPolity), level of inequality, and the incidence of civil conflict. For each of these variables, the means for the set of countries in SWIID are quite similar to the means for the set of survey countries. We have also examined the incidence of surveys in country-years that are experiencing conflict and country-years that are not, and we do not find significant biases: Surveys exist in 19.0 percent of country-years that are not experiencing conflict, and they exist in 15.3 percent of country-years that are experiencing conflict. Thus, although there are limits to the number of countries we can analyze using surveys, the sample of countries obtained using surveys seems reasonably unbiased with respect to the variables of central interest in the analysis here.

A second concern might be that the surveys do not adequately sample across groups. While we mitigate this concern to some extent by including only surveys that adequately cover the Fearon list of groups, there might still be biases in how many individuals from different groups are surveyed, particularly if some groups have members that are more difficult to survey than others. To assess this possibility, we calculated ELF – ethnolinguistic fractionalization – for each survey and then

[53]In this analysis, we focus on 88 countries since data on some key controls are missing for one of the countries in our dataset (Bosnia) and, therefore, it never enters our regressions.

Table B.14. Sample representativeness

	SWIID sample	Survey sample
Number of countries	136	88
Percentage of countries in:		
Central Europe	19.8	26.4
Latin America	16.2	12.5
Middle East	5.9	3.4
Africa	28.7	30.1
Neo-Europe	16.2	18.2
East Asia	8.8	5.7
South Asia	4.4	3.4
Average Real GDP/capita	\$9,836	\$10,288
Average F	.46	.50
Average P	.55	.58
Average xPolity	3.4	3.6
Average Gini (SWIID)	.38	.38
Percent of years with Prio25 civil conflict	.15	.17

Note. This table compares the sample of countries included in the dataset presented in this paper (88 countries) and the SWIID (137).

compared this with the ELF measures using Fearon's data. The correlation of the survey ELF and the Fearon ELF is an impressive .94, suggesting that countries that are included in our data set do a good job of capturing the distribution of individuals across groups.

A third potential concern with the surveys is that they inaccurately measure income, due to survey sampling strategies or to biases in how individuals report income on surveys. While this is a difficult issue to assess, we do find that the surveys and EPR data produce similar measures of group income per capita. Specifically, the correlation between the group GDP based on surveys and the group GDP based on the EPR data is .88.

Just as there are sources of concern with the survey data, there are also potential concerns with the main alternative, which is based on the spatial location of group members. One clear advantage with using the geo-coded data concerns its superior country coverage. The data set used by Cederman, Weidmann and Bormann (2015), for example, has data from 131 countries.

But there are a number of potential sources of bias using the geo-coded data. First, the spatial data share many of the same possible sampling biases that are a concern for the survey data. This is because the spatial approach must rely on expert estimates of the spatial location of group members, and thus risks measurement error because the experts themselves often do not have data on which to base their estimates of the group locations. The best data on which experts could draw would be some sort of careful survey or census, so any biases with respect to the coverage of groups in the survey data are going to be also present in the spatial data. Indeed, the biases are likely larger for the geo-coded data because the experts must estimate precisely where group members reside.

A second potential source of bias emerges from the need to assume either that particular geo-coded areas are occupied by only one group, or that individuals from different groups in the same geo-coded area have the same income. Neither assumption is attractive. There is substantial variation in the regional segregation of groups, and Morelli and Rohner (2013) link this segregation itself to civil conflict. And if one assumes that individuals from different groups occupy the same geo-coded area, one also has to assume that individuals from these different groups all have the same income – that is, to essentially assume what one is trying to measure.

A final concern stems from the treatment of urban areas. In some countries, groups might be relatively geographically segregated in the country side. But in urban areas, this is unlikely to be true, and it would be very challenging for country experts to accurately determine which ethnic groups are located in specific urban neighborhoods. Thus, urban areas, are typically excluded from group-based analyses using geo-coded data. With over half the world's population living in urban areas (Angel 2012), the fact that urban cells are discarded or measured inaccurately introduces a potential source of bias, particularly since an important potential source of within-group inequality (rural-urban inequality) is dismissed. Since, the urbanization of a country may be correlated with other factors that are related to civil war, the potential bias in the treatment caused by the treatment of urban areas may be correlated with conflict.

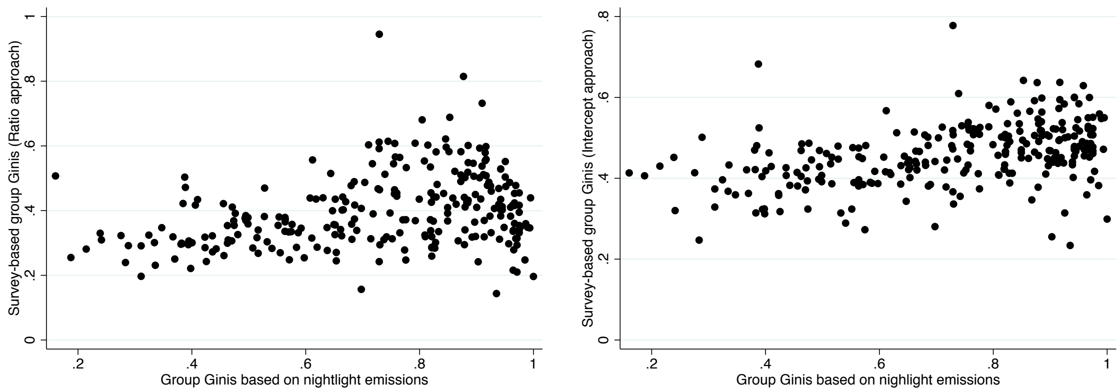
The problems with using the geo-coded data to measure the average income of groups are less severe than the problems associated with using the geo-coded data to measure within-group inequality, as in Kuhn and Weidmann (2015). Section 3.3 in the main text describes how KW calculate a Gini coefficient for each group based on variation across geographic cells. They call their measure of within-group inequality IGI.

As noted above, excluding urban groups will eliminate any within-group heterogeneity that is due to the rural-urban divide. In addition and very importantly, using spatial data to measure WGI will yield results that are sensitive to cell size, since the larger the size of the cell, the smaller the resulting within-group measure – in the limit, if the whole territory is assigned to one cell, within-group inequality would be zero. But the choice of cell size is arbitrary. It is not surprising, then, that Kuhn and Weidmann's (2015) own analysis shows a weak relationship between group Gini's based on nightlights and group Ginis based on (DHS) surveys, and that the relationship between these variables further weakens when urban areas are included.[54] And for measuring within-group inequality, the advantage of more countries is off-set to some extent by a greater difficulty in measuring the WGI of all groups. In KW, for example, measures are missing for nearly one-third of EPR groups. By comparison, the survey approach produces measures for 90 percent of the possible groups.

It is therefore useful to consider differences between measures using surveys and using night-lights. Although the two approaches produce similar measures of average group income, the same is not true for the measures of group income heterogeneity. Figure B.4 compares Gini coefficients computed using nightlight data (IGI) and surveys (G^I and G^R). The graph shows that there is a positive association between both set of measures, although divergences are substantial. The correlation between IGI and the survey-based Gini coefficients is a modest .41 and .34 for G^I and G^R respectively.

Above we described a number of reasons that could underline these differences. In addition, the geography of a group's homeland might lead to biases in the night-light based measures of within-group income heterogeneity. Table B.15, for example, examines the relation between the different inequality measures and the standard deviation of a group'w homeland elevations GROUP

[54]Their analysis also shows that when we compare a country-level Gini based on nightlights with a standard measure of the Gini, there is tremendous variation in the nightlight Gini at every level of the standard Gini; see Kuhn and Weidmann (2015), Figure 1.



(a) G^R vs. IGI

(b) G^I vs. IGI

Figure B.4. Survey-based versus Nightlight group-level Gini Coefficients

ELEV (SD). We might expect the diversity in the elevations occupied by a group to be correlated with diversity in nightlight readings: if travel is more difficult at higher elevations, individuals might tend to live more closely together in such regions (affecting the concentration of nightlight) but be no more or less well off than individuals living at lower elevations. The type of economic activity will also vary at different elevations.

Model 1 in Table B.15 shows the results from regressing IGI on GROUP ELEV (SD) – denoted in the table simply as ELEV to save space– along with country fixed effects. To facilitate interpretation, all variables in the table are standardized to have a mean of 0 and standard deviation of 1. We can see that the coefficient for ELEV is positive and large. A one standard deviation increase in ELEV is associated with a .24 standard deviation increase in IGI. Model 2 adds a control for the log of the group's GDP per capita and the result is essentially identical. Models 3 and 4 re-estimate models 1 and 2 using G^R as the dependent variable. The results are quite different, with the coefficient for G^R estimated very imprecisely.[55] Since there is data for more groups using nightlights than using surveys, models 3 and 4 have fewer observations than models 1 and 2. We therefore re-estimated models 1 and 2 using only the observation for which survey data is available. The results, in models 5 and 6, show that there is essentially zero difference with the estimates for

[55]We also estimated these models with G^I and obtained substantively similar results.

ELEV in models 1 and 2. Finally, since measures of horizontal inequality obviously treat information from all cells for a group identically (rather than looking for heterogeneity within groups), there is little reason to expect this same relationship between measures of horizontal inequality and ELEV(SD). To see if this is true, models 7 and 8 re-estimate models 1 and 2 using HI(LN) as the dependent variable. There is no relationship between this variable and ELEV. Finally, models 9 and 10 use country-level Gini coefficients (from Solt) and show that there is no correlation between country-level inequality and ELEV.

The results in Table B.15 suggest that the geo-coded measures of WGI could be biased upwards in places where variability in elevations is high. In fact, the correlation between the survey and the nightlight based WGI datasets increases considerably in places with low or moderate values of ELEV. The correlation between G^R (G^I) and IGI is .40 (.50) for ethnic homelands with a value of ELEV below the median and it goes down to .27 (.32) when values of ELEV above the median are considered. This analysis, along with the more general concerns discussed above regarding cell size and urban areas, lead us to conclude that considerable caution must be used when interpreting any results from analyses that use nightlights to measure WGI.

Table B.15. The relation between inequality measures and ethnic homeland's elevation

D.var.	IGI	IGI	G^R	G^R	IGI	IGI	HI(LN)	HI(LN)	CNTRY-GINI	CNTRY-GINI
ELEV	0.239***	0.243***	0.022	-0.015	0.226**	0.226**	-0.015	-0.001	0.010	0.008
	(0.000)	(0.000)	(0.829)	(0.898)	(0.019)	(0.019)	(0.762)	(0.984)	(0.515)	(0.583)
GDP_g		0.122		0.021				0.243***		0.024
		(0.133)		(0.843)				(0.000)		(0.629)
c	-0.543***	-0.186	-0.180	0.454	-0.090	-0.090	-0.015	0.701***	1.746***	0.028
	(0.000)	(0.471)	(0.305)	(0.285)	(0.581)	(0.581)	(0.786)	(0.000)	(0.000)	(0.569)
R^2	0.655	0.698	0.678	0.661	0.807	0.807	0.686	0.709	0.876	0.881
Obs	470	418	211	188	211	211	417	417	374	333

Note. This table regresses different inequality measures on the standard deviation of group's homeland elevation (denoted as ELEV to save space) and group's GDP (GDP_g, from G-Econ) . *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$

B.6. **Using EPR data to measure horizontal inequality.** This section describes how we use the EPR data to construct the measures used in section 6.2. In so doing, we compare our approach with CWB, the most recent paper to use EPR data to test horizontal inequality arguments.

The data set. The EPR data set we use is the 2014 update 2 version, and below we will refer to this data as "EPR", whereas CWB use an older version. CWB's data end in 2009, while the new EPR version extends to 2013. The post-2009 observations constitute about 16% of the data we use from EPR. In addition to different time periods, there exist substantive differences in the values of variables for the years covered by EPR and the CWB replication data. Consider the GECON measures of group well-being: there exist 752 observations (which represents about 7% of the observations in CWB) for which the CWB data have measures based on GECON but for which EPR has no information about the group's GDP based on GECON. Similarly, in EPR for the years up through 2009, there is GECON information about group GDP data for 1,129 group-years (which represents about 10% of the data during this pre-2009 time period) that are missing in the CWB data. And the control variables, while very similar, can also differ across CWB and EPR. For example, the correlation of EXCLUDED in CWB with EXCLUDED from EPR is .88.

Small groups and outliers. CWB advocate dropping from the regressions groups with a population of less that 500,000 out of concern for measurement error. We feel that while the concern may be valid, it is important to analyze the robustness of the results when no group is dropped, as we do in Table 4. This is particularly important since excluding small groups results in a substantial reduction in the size of the data set: for observations where the GECON data are available for a group, roughly one-third are for small groups in CWB. Small groups are also involved in a substantial number of conflict onsets, representing about one-third of the cases in CWB. But since concern about measurement error for such groups seems valid, it is alt important to analyze results where such groups are excluded. Our main point of departure with CWB in this regard is to note that if there are valid concerns about measurement error for small groups, this concern should influence the *construction* of horizontal inequality measures for large groups. In particular, failure to remove the smaller groups from the data set before creating the HI measures will create measurement error in the value of G, and thus will create measurement error for the HI variables for large groups. For this reason, when we estimate models excluding small groups, we also exclude such groups when the measures are created for the large groups. This can have a big impact on the horizontal inequality variables when the CWB definitions are used. The correlation of LOW constructed using all groups with LOW constructed excluding small groups is .58.[56] Similarly, CWB advocate omitting groups that are very large outliers from analysis. We follow this suggestion, but again we do so by excluding these groups before the measures are created.

Time varying measures. The GECON data vary over time, with different measures of group GDP in 1990, 1995, 2000 and 2005, and different measures of population in 1990, 2000 and 2010. Our measures of the horizontal inequality variables take advantage of the time-varying information, using the GDP and population data that is most proximate but previous to a given year.[57] This allows us to partially address concerns about reverse causality by lagging the measures of horizontal inequality. While the correlations of the variables constructed in different years are typically high, there are clear differences across years which could have an impact on the results, as could choices made in how to combine data across different years.

[56] It is useful to note that when constructing the corrected measures proposed in Appendix B.3, whether small groups are excluded in the data construction stage has little effect on the measures because such small groups exert almost no influence on $\bar{G}$. Indeed, the correlation of LOW CORRECTED constructed using all groups with LOW CORRECTED constructed excluding small groups is .99.

[57] So, for example, we use 1990 data for years 1990-94, and 1995 data for years 1995-1999.

Appendix C. Variable definitions and summary statistics

This section provides detailed definitions for the variables employed in the empirical analysis as well as a table of summary statistics.

C.1. **Variable definition.**

Conflict variables.

INCIDENCE: “Group level Armed conflict”. A binary measure taking a value of 1 for those years where an ethnic group is involved in armed conflict against the state resulting in more than 25 battle-related deaths. Ethnic groups are coded as engaged in conflict if a rebel organisation involved in the conflict expresses its political aims in the name of the group and a significant number of members of the group participate in the conflict. Source: Wucherpfennig et al. (2012).

INTENSITY: “Group level Conflict intensity”. We assign a value of 0 if group G is at peace in a given year, a value of 1 if there are events satisfying CONFLICT25_G and the total number of battle deaths that year does not exceed 1000, and a value of 2 if the number of battle deaths is larger than 1000. Source: Wucherpfennig et al. (2012).

ONSET: “Group level Conflict Onset”. A binary measure reflecting the first year in which a group enters a conflict, as defined in CONFLICT25_G above.

CONFLICT-SHARE: Share of years a group has been in conflict against resulting in more than 25 battle related deaths in the period 1992- 2010.

BATTLE DEATHS (BEST): Number of battle related deaths according to the *best* estimate from Lacina and Gleditsch (2005). One is added if the number of battle deaths is zero). In situations where the best estimate was missing, we used the low estimate instead.

BATTLE DEATHS (LOW): Log of the number of battle related deaths according to the *low* estimate from Lacina and Gleditsch (2005). One is added if the number of battle deaths is zero).

PEACEYEARS: Number of years since the last conflict observation. Source: Wucherpfennig et al. (2012).

Inequality measures.

G^R: Group Gini coefficient, computed using survey data and adjusted using the Ratio approach, as described in Section 3.1 and Appendix B.2. All available observations for a group are averaged and assigned to all the years in the period 1992–2010.

G^I: Group Gini coefficient, computed using survey data and adjusted using the Intercept approach, as described in Section 3.1 and Appendix B.2. All available observations for a group are averaged and assigned to all the years in the period 1992–2010.

G^U: Group Gini coefficient, computed using survey data, unadjusted. All available observations for a group are averaged and assigned to all the years in the period 1992–2010.

G_t^R-PRE: Group Gini coefficient, computed using survey data and adjusted using the Ratio approach. The value at time t of this variable is computed by averaging all available surveys in year t and in prior years. If no observations are available until period t, this variable is set to missing.

IGI: Group-level Gini coefficient computed using nighlight emissions. Source: Kuhn and Weidmann (2015).

LINEQ2: Measure of horizontal inequality, defined as $(log\frac{G_g}{G})^2$, where G_g is group's GDP per capita and G is the (unweighted) average of GDP per capita of all groups. Source: we have computed this measure using the surveys, GECON and nightlights data. Details are provided in the corresponding tables.

$\overline{\text{LINEQ2}}$: Measure of horizontal inequality, defined as $(log\frac{G_g}{\overline{G_g}})^2$, where G_g is group's GDP per capita and $\overline{G_g}$ is GDP per capita of non-g members.

HI(ABS): Group-level measure of horizontal inequality defined as $\text{HI(ABS)} = |g - \bar{G}|$, where g is group's GDP per capita and $\bar{G}$ is the weighted average (by group size) of the GDP per capita for all groups *other than* g. Source: we have computed this measure using the surveys, GECON and nightlights data. Details are provided in the corresponding tables.

HI(LN): natural log of HI(ABS). Source: we have computed this measure using the surveys, GECON and nightlights data. Details are provided in the corresponding tables.

LOW: It is defined as $max\{1, \frac{G_g}{G}\}$, where G_g is group g per capita GDP and G is the average of the per capita GDP of all groups.

$\overline{\text{LOW}}$: It is defined as $max\{1, \frac{G_g}{\overline{G_g}}\}$, where G_g is group g per capita GDP and $\overline{G_g}$ is the per capita GDP of non-g members.

HIGH: It is defined as $max\{1, \frac{G}{G_g}\}$, where G_g is group g per capita GDP and G is the average of the per capita GDPs of all groups.

$\overline{\text{HIGH}}$: It is defined as $max\{1, \frac{\overline{G_g}}{G_g}\}$, where G_g is group g per capita GDP and $\overline{G_g}$ is the per capita GDP of non-g members.

Controls.

GDP: log of real GDP per capita, lagged one year. The source is the Penn World Tables (2015).

POP: log of the population in millions, lagged one year, as reported by the Penn World Tables (2015).

XPOLITY: democracy score based on Polity IV, lagged one year. It combines 3 out of the 5 components of Polity IV (XCONST, XRCOMP, XROPEN) and leaves out the two components (PARCOMP and PARREG) that are related to political violence, and hence are likely to be endogeneous. It ranges from -6 (maximum level of autocracy) to 7 (maximum level of democracy). See Vreeland (2008) for details.

GROUP GDP: Survey-based group GDP per capita, lagged one year.

POVERTY2: Percentage of the total population with income lower than 2 dollars a day. Source: World Bank.

INFANT MORTALITY: Infant mortality rate. This variable is a snapshot from the year 2000. Source: PRIOGRID, accessed through Growup portal, `https://growup.ethz.ch/`.

EXCLUDED: dummy variable indicating whether the group is excluded from power. Source: GrowUp portal, `https://growup.ethz.ch/`.

N. EXCLUDED GROUPS: Number of groups in the country excluded from power. Source: GrowUp portal, `https://growup.ethz.ch/`.

GROUP DIAMONDS: dummy variable indicating whether the group has diamonds in its homeland. Source: PRIOGRID, through the Growup portal, `https://growup.ethz.ch/`.

GROUP OIL: dummy variable indicating whether the group has oil in its homeland. Source: PRIOGRID, through the Growup portal, `https://growup.ethz.ch/`.

GROUP ELEV. (SD): the standard deviation of the elevation of the ethnic homeland. Source: GrowUp portal, `https://growup.ethz.ch/`.

POW. BALANCE. Demographic power balance between the group and the group(s) in power. Denoting the populations of the group and the group(s) in power as s and S, respectively, the power balance is defined as s/(s+S) if the group is excluded, and as s/S otherwise.

GIP: dummy variable indicating whether the group has access to power. Source: GrowUp portal, urlhttps://growup.ethz.ch/

EXCLUDED: dummy variable indicating whether the group is excluded from power in year $t-1$(i.e., it is defined as 1-GIP. Source: GrowUp portal, urlhttps://growup.ethz.ch/.

GROUP SIZE: Relative size of the group. Source: GrowUp portal, urlhttps://growup.ethz.ch/.

LINGUISTIC FRAC : Within-group Linguistic fractionalization index. Source: Growup portal, `https://growup.ethz.ch/`.

RELIGIOUS FRAC : Within-group religious fractionalization index. Source: Growup portal, `https://growup.ethz.ch/`.

REGIONAL : A dummy measuring whether a group is regionally concentrated. Source: Growup portal, `https://growup.ethz.ch/`.

URBANIZED: The proportion of a group?s homeland that is urbanized. Source: Prio-Grid.

RAINFALL : Average annual rainfall in the areas controlled by a group. Source: Prio-Grid.

AGRICULTURE: Percentage of a group?s area that is covered with agricultural production. Source: Prio-Grid.

C.2. **Summary statistics.** Tables C.16 and C.17 provide summary statistics for the variables employed in Sections 4 and 6.2, respectively.

Table C.16.

variable	**Obs.**	**mean**	**sd**	**min**	**max**
INTENSITY	5945	0.04	0.21	0.00	2.00
BATTLE DEATHS (BEST)	5945	0.17	0.94	0.00	9.74
BATTLE DEATHS (LOW)	594	0.15	0.83	0.00	9.21
G^R	7369	0.39	0.12	0.05	0.95
G^I	7369	0.45	0.09	0.05	0.78
G^u	7369	0.29	0.08	0.05	0.66
G^R-PRE	5045	0.39	0.13	-0.00	0.95
HI(LN)	6942	5.79	1.90	-2.43	10.86
HI(LN)-PRE	5045	6.36	1.93	-6.76	13.42
LINEQ2(SURVEY)	6942	0.08	0.21	0.00	2.28
LINEQ2(GECON)	4578	0.14	0.30	0.00	3.47
LOW(SURVEY)	7369	0.70	0.59	0.00	3.54
HIGH(SURVEY)	7369	0.54	0.57	0.00	4.53
LOW(GECON)	4578	1.23	0.46	1.00	6.44
LOW(GECON)	4578	1.14	0.30	1.00	3.15
EXCLUDED	5418	0.43	0.50	0.00	1.00
GROUP SIZE	7369	0.15	0.19	0.01	0.96
ELEV(SD)	4724	282.77	246.01	5.77	1617.65
GROUP DIAMONDS	7392	0.17	0.38	0.00	1.00
GROUP OIL	7392	0.22	0.42	0.00	1.00
POP	5847	2.88	1.49	-0.52	7.10
XPOLITY	6824	2.85	4.03	-5.00	7.00
GDP	5847	8.20	1.21	5.92	10.79

Notes. This table presents summary statistics of the variables considered in Section 4.

Table C.17.

variable	**Obs.**	**mean**	**sd**	**min**	**max**
ONSET	5691	0.01	0.08	0.00	1.00
HI(LN)	4545	-2.87	1.63	-9.84	0.92
LINEQ2	4578	0.14	0.30	0.00	3.47
LOW	4578	1.23	0.46	1.00	6.44
HIGH	4578	1.14	0.30	1.00	3.15
$\overline{\text{LOW}}$	4541	1.27	0.56	1.00	8.88
$\overline{\text{HIGH}}$	4535	1.26	0.56	1.00	8.88
HI(LN), SMALL	3550	-2.99	1.62	-9.84	0.85
LINEQ2, SMALL	3583	0.11	0.28	0.00	3.47
LOW, SMALL	3583	1.17	0.39	1.00	6.44
HIGH, SMALL	3583	1.14	0.31	1.00	3.15
$\overline{\text{LOW}}$, SMALL	3431	1.26	0.57	1.00	8.88
$\overline{\text{HIGH}}$, SMALL	3425	1.24	0.57	1.00	8.88
EXCLUDED	15816	0.58	0.49	0.00	1.00
DOWNGRADED	16579	0.02	0.15	0.00	1.00
GROUP SIZE	17682	0.19	0.28	0.00	1.00
POSTWAR	19027	0.28	0.72	0.00	6.00
ONGOING CONFLICT	18200	0.24	0.43	0.00	1.00
GDP	17836	8.66	1.16	5.09	11.88
POP	17836	3.14	1.85	-0.87	7.21

Notes. This table presents summary statistics of the variables considered in Section 6.2. All data comes from the Growup portal.