Fragmentation or Inequality? Ethnic Divisions and Conflict in Sub-Saharan Africa

Boris Gershman* Ameesh Upadhyay†
American University Knox College

April 2025

Abstract

This paper examines the relationship between local ethnic divisions and conflict in Sub-Saharan Africa. Using census subsamples and large-scale household surveys, we construct a new subnational dataset on ethnic inequality capturing group-level differences in education, asset ownership, and access to basic amenities for several hundred regions in thirty-five countries. To distinguish between deep-rooted and more recent ethnic divisions, we incorporate groups from our sample into *Ethnologue*'s linguistic tree model and generate alternative measures of both ethnic fragmentation and inequality based on ancestral languages. Our analysis, leveraging within-country variation and accounting for numerous regional characteristics, reveals a robust positive relationship between ethnic fractionalization and conflict, especially when using deeper linguistic cleavages to define distinct groups. In contrast, ethnic inequality shows no systematic association with conflict frequency or severity. These findings suggest the primacy of ethnic identity over socioeconomic disparities between groups as a driver of local conflict.

Keywords: conflict, ethnolinguistic diversity, ethnic inequality, Sub-Saharan Africa, subnational analysis

JEL Classification Numbers: D31, D74, O10, O15, Z13

*Department of Economics, American University, 4400 Massachusetts Avenue NW, Washington, DC 20016-8029 (e-mail: boris.gershman@american.edu).

†Department of Economics, Knox College, 2 East South Street, Galesburg, IL 61401 (e-mail: aupadhyay@knox.edu).

1 Introduction

Population diversity has long been viewed as an important determinant of economic growth and development. A vast empirical literature explores many dimensions of societal differences including economic inequality, religious, cultural, and genetic diversity, with a special focus on ethnolinguistic divisions in the context of Sub-Saharan Africa.[1] In principle, population diversity can have both costs and benefits, so its net effect on socioeconomic outcomes is theoretically unclear.[2] Given the crucial role of lasting peace in securing economic prosperity, arguably the most important social cost of population diversity is its potential to foster conflict (Arbatlı et al., 2020). This issue is particularly relevant for Africa, a diverse continent consistently displaying high levels of organized violence (Davies et al., 2024).

Despite the growing number of studies linking ethnic divisions to conflict, the overall evidence on the nature of this relationship remains mixed. To some extent, this reflects the multitude of approaches to measuring the relevant aspects of diversity. For example, while earlier studies have focused on the indices of ethnolinguistic fractionalization and polarization, recent work has argued that between-group ("horizontal") ethnic inequality represents a more salient determinant of conflict (Hillesund et al., 2018). Another point of contention has been the choice of relevant group identities and the importance of accounting for cultural proximity in diversity indices (Desmet et al., 2012). Furthermore, the literature has gradually moved away from cross-country to subnational analyses, posing the challenge of constructing reliable metrics at the local level (Gershman and Rivera, 2018).

In this paper, we contribute to the ongoing debate by building a new comprehensive dataset on various dimensions of local ethnolinguistic divisions in Sub-Saharan Africa and revisiting the relationship between diversity and conflict at the subnational level. Our dataset covers 35 countries, 391 first-level administrative regions, and 853 unique ethnolinguistic groups. We use microdata from large-scale surveys to produce ethnic inequality indices capturing group-level differences in education, asset ownership, and access to basic amenities, in addition to standard measures of fractionalization and polarization. Following a growing literature emphasizing the importance of cultural and linguistic relatedness among groups, we distinguish between deep-rooted and contemporary ethnolinguistic cleavages. To do so, we match all groups from our sample to the *Ethnologue* database and aggregate our data on population composition at different tiers of the global linguistic tree,

[1] See, for example, Cuesta and Wantchekon (2016) and Gershman and Rivera (2018).

[2] These trade-offs are reviewed in Alesina and La Ferrara (2005), Montalvo and Reynal-Querol (2014), and Ginsburgh and Weber (2020).

corresponding to more or less recent hypothesized ancestral languages. This results in a hierarchy of diversity indices, from those based on cleavages originating in the very distant past to those focusing solely on contemporary language distinctions.

We combine our new subnational diversity database with georeferenced data on conflict events and fatalities dating back to the late 1990s, along with other regional-level characteristics. Our regression analysis reveals a positive relationship between ethnolinguistic fractionalization (ELF) and conflict, which is stronger in magnitude and statistical significance for diversity metrics based on deep linguistic cleavages. After accounting for country fixed effects and numerous potential regional confounders, we find that, across specifications, a one-standard-deviation increase in deep-rooted fractionalization is associated with roughly a 20-35% rise in conflict events and an even larger percentage increase in fatalities. In contrast, we find that none of our ethnic inequality metrics are systematically related to conflict. Even in those cases where a significant association is found, it disappears after accounting for ELF. Similarly, there is no robust pattern linking conflict to polarization. Our results hold across model specifications and for a variety of conflict measures from alternative data sources.

Our main contribution is twofold. First, we advance the ongoing effort to leverage microdata from large-scale surveys for building standardized datasets on population diversity across countries and regions. Our new subnational measures of ethnic inequality offer a broad coverage of Sub-Saharan Africa based on high-quality sources, employ a variety of underlying socioeconomic indicators, and account for ancestral relationships between ethnolinguistic groups. Our general approach to data construction is closely related to Gershman and Rivera (2018) but additionally covers measures of ethnic inequality.[3]

The units of analysis in our study are the first-level administrative regions, which naturally form politically and economically relevant geographic entities. This allows us to focus on within-country variation in diversity and conflict while accounting for nationwide fixed effects. Since we cover a fairly standard set of regions in their current or recent boundaries, our new data are compatible with most existing subnational-level databases and represent a ready-to-use source for future research on diversity.

Our survey-based approach to measuring population diversity utilizes information on actual individuals and households and differs from the increasingly popular GIS-based

[3]Fjelde and Østby (2014) is a notable earlier contribution using survey data (the Demographic and Health Surveys) to measure subnational ethnic inequality in Africa. However, the authors focus exclusively on the difference between a region's largest ethnic group and the rest of its population. In contrast, we consider all groups residing in a region and offer a more comprehensive coverage of Africa's linguistic diversity by using superior data sources and accounting for relatedness between groups.

analyses that overlay digital maps of population counts, ethnic homelands, and measures of material wealth around the world. Gershman and Rivera (2020) compare these two approaches and highlight the sources of measurement errors associated with the use of GIS. Most importantly, even the best available maps of ethnic or linguistic homelands are inaccurate since they do not account for recent migrations and are generally not designed to capture the cases of overlapping groups, especially in urban areas. As a result, GIS-based metrics of subnational fractionalization are only moderately correlated with the higher-quality metrics based on censuses and large-scale population surveys. In the context of ethnic inequality, the scope for measurement error is even greater since one additionally needs fine-grained information on material well-being at the ethnic-group level, typically derived from night lights data (Kuhn and Weidmann, 2015; Alesina et al., 2016; Leipziger, 2023). The latter are merely an imperfect proxy that has been argued to be particularly poor in small and rural areas (Gibson et al., 2020). Furthermore, since this approach relies on a single night lights measure, it cannot distinguish between different dimensions of ethnic inequality. Overall, despite its appealing simplicity, the use of GIS methods for measuring ethnic inequality, particularly at the subnational level, is associated with multiple layers of measurement error.

Second, we contribute to the large empirical literature linking ethnolinguistic diversity and conflict. While early cross-country studies (Fearon and Laitin, 2003; Collier and Hoeffler, 2004) failed to identify a significant link between ELF and conflict and emphasized the role of polarization (Montalvo and Reynal-Querol, 2005), later, more theoretically-grounded work found both measures to be positively related to conflict (Esteban et al., 2012). Furthermore, Desmet et al. (2012) showed that a positive country-level relationship between ELF and conflict becomes more apparent for measures based on deep-rooted ethnolinguistic cleavages. Our analysis is performed at the subnational level and focuses on Sub-Saharan Africa, but is largely in line with this more recent cross-country evidence. It also complements the finding on the negative association between deep-rooted diversity and local public goods provision (Gershman and Rivera, 2018). Conceptually, the observed pattern is consistent with the intuition that ethnolinguistic divisions, especially those originating in the distant past, give rise to coordination problems, mistrust, and inter-group animosity that spur conflict.

Our null findings for ethnic inequality contrast the research arguing that it is the overlap of ethnic identity markers and socioeconomic inequality, rather than diversity per se, that drives conflict (Stewart, 2008). Evidence supporting this view was reported in several national and subnational studies (Cederman et al., 2011; Fjelde and Østby, 2014; Houle

and Bodea, 2017), although these analyses all rely on different metrics of both horizontal inequalities and conflict.[4] On the other hand, some studies, like ours, failed to find a significant overall relationship (Østby, 2008; Huber and Mayoral, 2019) or only found it for a specific dimension of ethnic inequality such as differences in education levels (Alcorta et al., 2018). We add to this mixed set of results by providing new evidence based on high-quality subnational data, capturing different types of ethnic inequality across regions and accounting for linguistic relatedness between groups.

Overall, our findings suggest that the fragmentation of population into many ethnolinguistic groups, particularly when they are more culturally distinct, is a better predictor of conflict than socioeconomic inequalities between them. Thus, identity-based factors appear to be more salient than issues associated with group differences in socioeconomic status. This result resonates with a recent finding that ethnolinguistic identity per se, rather than relative income, is the primarily driver of regional support for secession (Desmet et al., 2025).

Generally, while there are plausible mechanisms potentially linking ethnic inequality to conflict (Hillesund et al., 2018), this relationship is theoretically ambiguous. Indeed, some research has argued that economically similar groups may be *more* likely to fight, and ethnicity- rather than class-based conflict is more salient in the presence of economic inequality (Esteban and Ray, 2008; Ray and Esteban, 2017). A poorer group may simply lack the resources to rebel or resists and an improvement in its relative standing may enhance its chances of success in conflict. While frustration and resentment stemming from group differences in socioeconomic outcomes can provide motivation, successful mobilization for collective action also requires organization and resources. Given this conceptual ambiguity and the presence of multiple countervailing mechanisms, our null result on ethnic inequality is not particularly surprising.[5]

The rest of the paper is organized as follows. Section 2 introduces the new dataset. Section 3 presents our main results and robustness tests. Section 4 concludes. Further details on data construction, definitions of all variables, and additional analyses are collected in the appendices.

[4]Relatedly, Baldwin and Huber (2010) and Alesina et al. (2016) argued that ethnic inequality is a better predictor of public goods provision and overall development at the country level than ELF.

[5]An interesting line of research suggests that perceived inequality and group status do not necessarily correspond well to the actual group-level differences and are subject to manipulation (Hillesund et al., 2018). In this case, objective ethnic inequality measures may fail to capture relevant grievances.

2 Data

2.1 Ethnolinguistic groups and socioeconomic indicators

Our new dataset on local population diversity covers 391 first-level subnational administrative regions across 35 countries in Sub-Saharan Africa.[6] For each country, we selected a primary source of microdata based on three main priorities. First, we preferred surveys containing more detailed listings of ethnolinguistic groups. Second, given our goal of measuring regional ethnic inequality, each chosen survey had to contain relevant socioeconomic indicators for individuals or households. Third, we aimed to maximize sample size and representativeness at the subnational level. Overall, our microdata come from three sources: census subsamples from the Integrated Public Use Microdata Series (IPUMS), Demographic and Health Surveys (DHS), and Multiple Indicator Cluster Surveys (MICS). Both the average and the median year of observation for population diversity metrics are around 2010, with some variation dictated by data quality and availability constraints. The countries in our sample, along with respective primary sources of microdata, are listed in the first two columns of appendix table B.1.

We follow the methodology of Gershman and Rivera (2018) and standardize the notion of a unique ethnolinguistic group by matching the self-reported identities from original surveys to corresponding spoken language codes in the *Ethnologue* database (Lewis, 2009). This process ensures comparability and weeds out arbitrarily defined categories.[7] In particular, we dropped cases that did not match any *Ethnologue* language code and those matched to non-native languages, resulting in a total of 853 unique ethnolinguistic groups. In contrast to some earlier research that focused only on "politically relevant" ethnic groups (Posner, 2004; Cederman et al., 2011), we placed no such restrictions.[8] An additional benefit of our matching process is that it incorporates the groups in our dataset into *Ethnologue*'s language tree model, which captures the ancestral relationships between languages and allows us to distinguish between recent and deep-rooted ethnolinguistic cleavages when constructing our diversity measures, as described in section 2.2.

[6] We use the same set of regional boundaries as Gershman and Rivera (2018). See figure 2 for a map.

[7] For example, the 2013-14 DHS dataset for the Democratic Republic of the Congo delineates "ethnic groups" based on geographic locations making this identifier unusable.

[8] As discussed by Hillesund et al. (2018), a focus on politically relevant groups in the study of conflict is conceptually problematic. First, this approach excludes not-yet-mobilized groups that may nonetheless be important for assessing the risk of conflict. Second, the coding of "relevant" groups is often post-hoc, that is, based on observed conflict events, which introduces bias in favor of finding a significant pattern. Our approach avoids these potential pitfalls.

We faced the following trade-off when choosing between IPUMS and DHS/MICS as microdata sources. On the one hand, IPUMS data typically provide both much larger population samples and more detailed listings of ethnolinguistic groups (if this information was collected in respective censuses). On the other hand, DHS/MICS tend to report a wider range of socioeconomic outcomes that can be used to construct a larger variety of ethnic inequality metrics. Given our priorities, we relied on IPUMS as the primary data source for 14 countries in our sample, with the remaining 21 covered by DHS or MICS. Whenever IPUMS did not report some of the socioeconomic outcomes of interest, we drew upon DHS/MICS alternatives to compute corresponding ethnic inequality indices. For example, some of the IPUMS surveys report only a small set of asset ownership indicators, in which cases we employ DHS/MICS. We refer to these auxiliary data sources as "secondary" and list them in the third column of appendix table B.1.

As column 4 of the same table makes clear, our primary and secondary data sources sometimes provide a substantially different classification of ethnolinguistic groups. To take a couple of extreme examples, for Benin and Ghana, DHS only list 8, while IPUMS data cover around 40 distinct groups in each case. The main reason for this discrepancy is that DHS surveys occasionally aggregate multiple related smaller groups into a single ethnolinguistic cluster. Although one has to be mindful of this issue, it is not a major concern since the vast majority of our indicators are well-covered by primary data sources and we focus on those in the main text.[9]

We aimed to capture a variety of socioeconomic divisions between groups. As reviewed by Leipziger (2023), there is no universally accepted "best" measure of horizontal ethnic inequality, and previous research looked at its various economic and social dimensions. The former capture income and wealth that are often associated with the ownership and consumption of private goods, whereas the latter include education and health that typically reflect access to locally provided public goods in a developing economy setting. Exploiting the availability of multiple outcomes in our survey data, we construct different ethnic inequality metrics without making a priori assumptions regarding the most prominent sources of societal divisions.[10] Specifically, following the above classification, our "economic" indicators include binary measures of asset ownership (television, car, radio, flush toilet), whereas the "social" indicators are represented by educational attainment (in years). In

[9]The last row of appendix table D.2 reports the number of subnational regions covered by the primary data sources for each diversity index.

[10]Alcorta et al. (2018) argue that grievances triggered by educational inequality are more likely to incentivize collective action, while differences in asset ownership may have the opposite effect by limiting the ability of disadvantaged groups to rebel.

addition, we look at household access to electricity and high-quality water sources.[11] Indicators that are measured on a categorical scale in the original surveys (such as source of water, type of latrine, or flooring material) are converted into binary variables using the "high quality" standard from Smits and Steendijk (2015). For example, we consider the binary measure of access to flush toilet (contrasted with pit latrines or other lower-quality options). Detailed definitions of the underlying socioeconomic outcomes are available in appendix A.

Our approach is different from attempts to aggregate the information on socioeconomic characteristics and build indices of overall individual or household wealth using principal component analysis or similar techniques (McKenzie, 2005; Smits and Steendijk, 2015). We are mainly motivated by the conceptual differences in outcomes outlined in the previous paragraph, but aggregate indices have other important limitations. For example, not all surveys report the same outcomes and classification of ethnolinguistic groups, so a uniform measure of average group-level "well-being" is hard to achieve. Even when there is sufficient coverage, the choice of appropriate weighting scheme for different index components is highly non-trivial, especially with discrete data (McKenzie, 2005; Howe et al., 2008). Furthermore, the extent of agreement between aggregate indices and conventional measures of socio-economic position remains unclear (Howe et al., 2008; Filmer and Scott, 2012).

2.2 Measuring diversity

We focus on three popular diversity measures: fractionalization, polarization, and between-group ethnic inequality. The index of ethnolinguistic fractionalization reflects the degree to which a region is comprised of many small groups. It is calculated as follows:

$$\text{ELF} = 1 - \sum_{i=1}^{N} s_i^2,$$

where s_i is the population share of group $i \in \{1, ..., N\}$ and N is the total number of groups. This value can be interpreted as the probability that two randomly chosen individuals from a given region belong to different ethnolinguistic groups. Fractionalization is highest when each resident of the region has a different ethnicity.

[11] We also examined but ultimately dropped the indicators of refrigerator, motorcycle, and bicycle ownership, primary and secondary school attendance, basic literacy, underweight status of children, and share of home births. These outcomes were either highly correlated with other indicators on our list or had poor availability (or both).

As mentioned in the introduction, some research has argued that ethnolinguistic polarization (ELP) is a more relevant predictor of conflict than ELF. Following Montalvo and Reynal-Querol (2005), we calculate the ELP index as follows:

$$\text{ELP} = 4\sum_{i=1}^{N} s_i^2(1 - s_i).$$

Unlike ELF, it captures the extent to which population is divided into large groups and reaches the maximum of 1 when a region is evenly split between two ethnicities.

Both ELF and ELP indices are based solely on the population shares of ethnic groups and incorporate no information on socioeconomic differences between them. The presence of such divisions is captured by ethnic inequality metrics. Following Baldwin and Huber (2010), we employ the between-group Gini index which may be calculated for any relevant socioeconomic outcome w. Suppose the mean outcomes for a pair of groups i and j are given by $\bar{w}_i$ and $\bar{w}_j$, respectively, and let $\bar{w}$ be the mean outcome for the entire region. Then, between-group inequality (BGI) index is defined as

$$\text{BGI} = \frac{1}{2\bar{w}} \sum_{i=1}^{N} \sum_{j=1}^{N} s_i s_j |\bar{w}_i - \bar{w}_j|.$$

The resulting measure may be interpreted as the (normalized) expected absolute difference in values of outcome w between two randomly chosen residents of a given region, assuming w is evenly distributed *within* each group. As explained in the previous section, we use a variety of underlying socioeconomic indicators available in our survey microdata to build corresponding BGI indices.

For most outcomes, population shares and group averages are calculated from household-level data, and ethnolinguistic groups are defined based on self-reported identities of household heads. For years of education, BGI is computed using individual-level data on adults aged 15-49 and respective ethnicity (or, in some cases, native language) identifiers. Both ELF and ELP indices are also calculated from individual-level data.[12]

The indices introduced so far treat all groups as equally distinct from an ethnolinguistic perspective. However, in reality, groups may be more or less linguistically and culturally distinct depending on the extent of shared ancestry. This may naturally affect their mutual interactions. To account for this nuance, previous studies have incorporated the information on linguistic relatedness in standard diversity measures. Here, we implement the method used by Desmet et al. (2012) and Gershman and Rivera (2018) in the

[12]MICS only provide the information on ethnicity and/or native language of household heads and we assume the same identity for other members of the same household.

national and subnational context, respectively. Specifically, we use *Ethnologue*'s linguistic tree model to aggregate basic ethnolinguistic groups to the levels of their more and less recent hypothesized ancestral language groups and compute diversity indices at each such level of aggregation $k = 1, \ldots, 13$.[13] The most disaggregated level, $k = 13$, corresponds to contemporary ethnolinguistic groups and their languages, that is, the standardized identities reported in our sample. In contrast, at the most aggregated level, $k = 1$, all ethnicities are grouped into just six major language families of the African continent: Afro-Asiatic, Niger-Congo, Nilo-Saharan, Creole, Indo-European, and Khoisan.[14] Higher aggregation levels (lower values of k) correspond to deep-rooted cleavages that emerged in the distant past, whereas standard diversity measures for $k = 13$ reflect contemporary language distinctions.

This procedure yields 13 versions of each diversity index. We denote them as ELF(k), ELP(k), and BGI(k) referring to the level of linguistic aggregation in parentheses.[15] While subnational ELF(k) and ELP(k) metrics for Sub-Saharan Africa were calculated and explored by Gershman and Rivera (2018), to the best of our knowledge, the present paper is the first to introduce subnational BGI(k) indices.[16] Hence, in the rest of this section, we focus on this part of our local diversity dataset. More specifically, we describe four groups of BGI indices based on educational attainment, access to electricity, ownership of television and car, which fairly represent the whole spectrum of ethnic inequality metrics in our dataset. The former two indices largely reflect inequality in access to local public goods, whereas the latter are based on the ownership of private household durables. According to a classification discussed earlier, education-based BGI represents the "social" dimension of ethnic inequality, while the other metrics focus on its "economic" dimension.

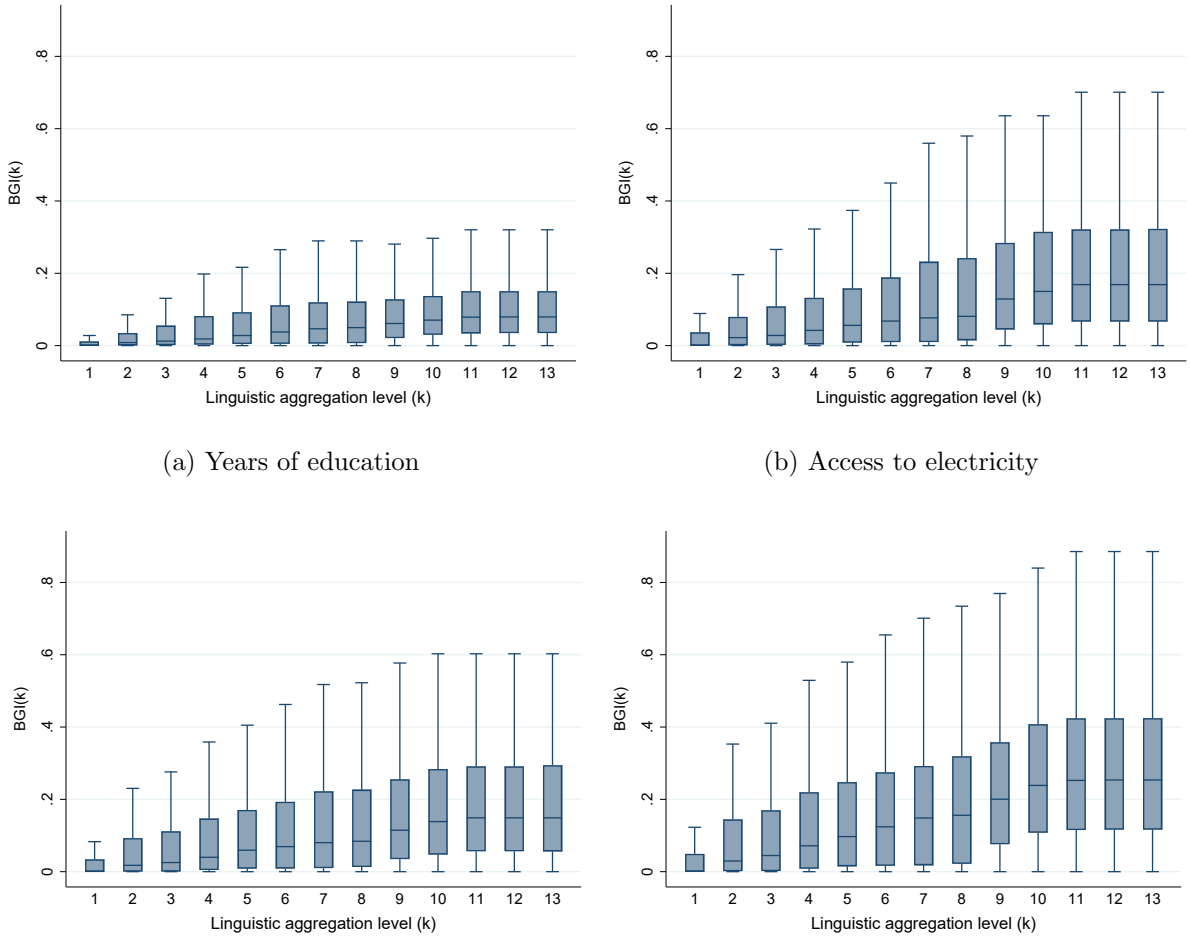
Figure 1 summarizes the distributions of BGI(k) indices for the four underlying indicators in the form of box-and-whiskers plots, sorted on the horizontal axis by the level of

[13]The presence of 13 aggregation tiers is dictated by *Ethnologue*'s delineation of ancestral groups for languages in our sample.

[14]The latter three families are very small and largely represent several "mixed" languages of West Africa (Creole) and indigenous South African groups (Khoisan). The only local Indo-European language is Afrikaans, and we also include small English-speaking groups native to South Africa, Botswana, and Namibia.

[15]Appendix C describes the process of constructing BGI(k) indices in detail using a specific region from our sample. Selected summary statistics for all diversity indices are reported in appendix table D.2.

[16]Hodler et al. (2020) develop an index of "ethnic stratification" that represents an alternative way to incorporate the information on linguistic relatedness when measuring ethnic inequality. Using the Afrobarometer data, they show that ethnic stratification measured at the town or village level is negatively related to trust.



(a) Years of education

(b) Access to electricity

(c) Television ownership

(d) Car ownership

Figure 1: Box-and-whiskers plots for selected BGI(k) indices.

Notes. Box-and-whiskers plots present the following characteristics of the respective distributions: lower quartile (bottom of the box), median (solid line segment), upper quartile (top of the box) and the lower and upper adjacent values.

linguistic aggregation. Generally, BGI indices are lower for smaller k, mechanically reflecting the reduction in the number of distinct groups at upper tiers of the linguistic tree.[17] To take the most extreme example, for over a third of regions in our sample, all contemporary ethnic groups belong to a single major language family, that is, BGI(1) is trivially equal to zero. However, for larger k, we observe substantial variation in ethnic inequality indices.

Interestingly, the distributions of BGI(k) for educational attainment are more compressed compared to other indicators. On the other hand, BGI for car ownership has both

[17]Appendix C contains a formal proof for a simplified case that BGI(k) is non-decreasing in k.

the widest range and the highest median. In general, we see a higher degree of inequality in the ownership of assets that may be considered luxurious in a low-income country setting, such as car, flush toilet, and finished flooring, as opposed to more accessible television and radio. For example, average BGI(13) index for the former three categories is close to 0.3, whereas it is 0.2 for television and just 0.07 for radio (see appendix table D.2).

Figure 2 presents the spatial distributions of BGI(k) in electricity access for $k = 1$, 5, 9, 13. Consistent with earlier observations, the map is lighter overall (showing lower ethnic inequality) at higher aggregation levels, when the number of distinct groups per region is smaller. Nevertheless, even at a relatively high level of aggregation $k = 5$, there is a lot of variation across regions, despite the fact that numerous language groups (such as all Bantu speakers) are "merged" into one. Generally, the observed variation in BGI captures both population fragmentation into groups and their differences in electricity access.

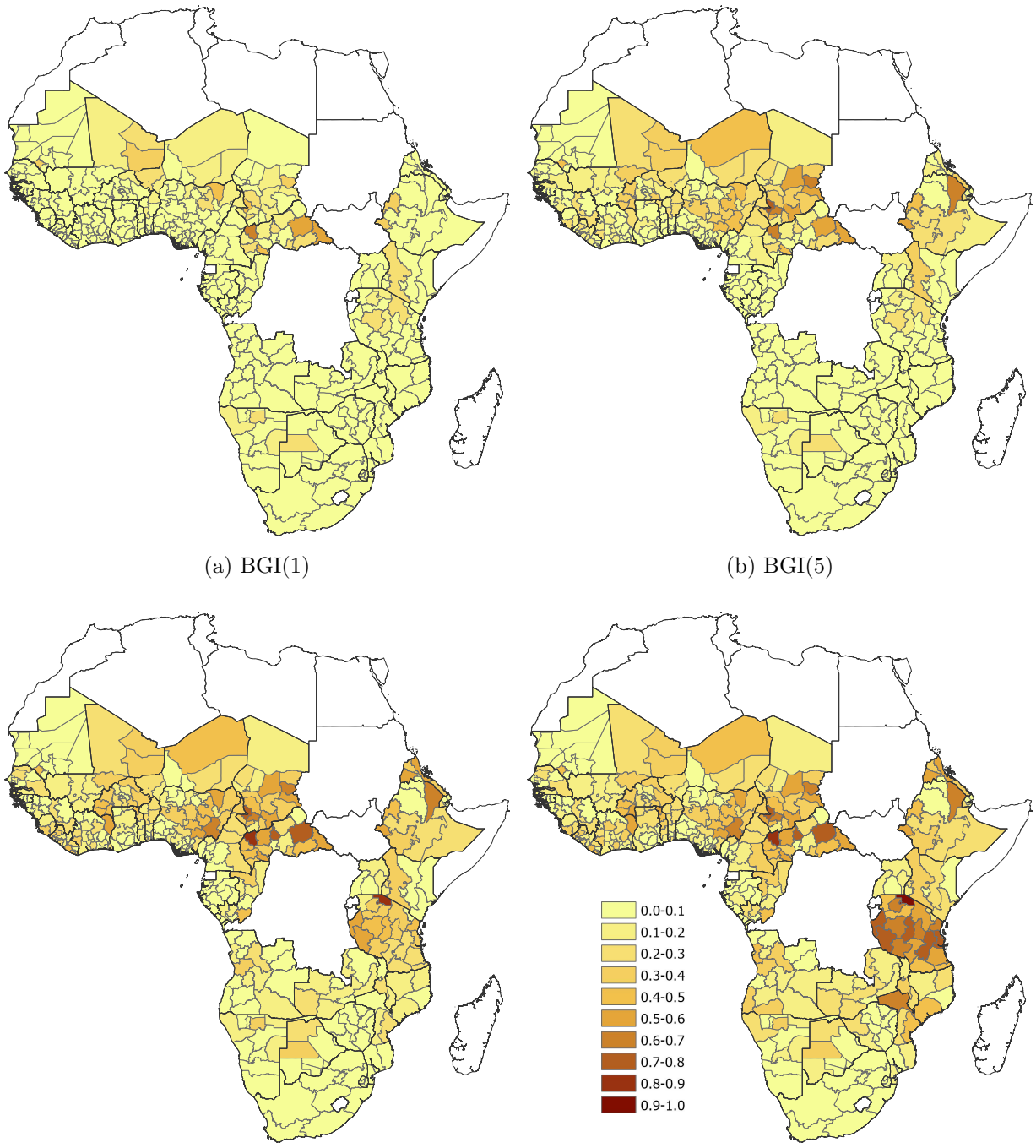
Regions with higher ethnic inequality in one outcome (e.g., electricity access) are also typically more unequal in other dimensions. As shown in the appendix table D.3, pairwise correlation coefficients for BGI(k) in access to electricity and years of education are in the 0.2–0.6 range, with roughly similar interval for ethnic inequality in car ownership. The correlation is stronger for the indicator of television ownership, falling within the 0.4–0.8 range, which is intuitive, given the dependence of TV usage on electricity access.[18] Not surprisingly, we typically observe the highest pairwise correlations between BGI indices measured at the same level of linguistic aggregation since the number of distinct groups is identical at a given tier k.

Ethnic inequality measures are moderately positively correlated with ELF at the corresponding tiers, as illustrated in the appendix figure D.1. The pairwise correlation coefficients between ELF and our baseline set of BGI indices fall roughly within the 0.4–0.7 range. The association between BGI and ELP is generally weaker compared to ELF.

2.3 Measuring conflict

Our primary source of georeferenced data on conflict events and fatalities is the Armed Conflict Location and Event Database (Raleigh et al., 2010), or ACLED, which has been widely used in the literature. ACLED captures a broad range of violent and non-violent actions involving political agents, including governments, rebels, militias, identity groups, political parties, external forces, rioters, protesters, and civilians. It is derived from numerous local, national, and international sources including press accounts, books, research

[18]The numbers are similar for access to high-quality water sources, presence of flush toilet, and radio ownership.

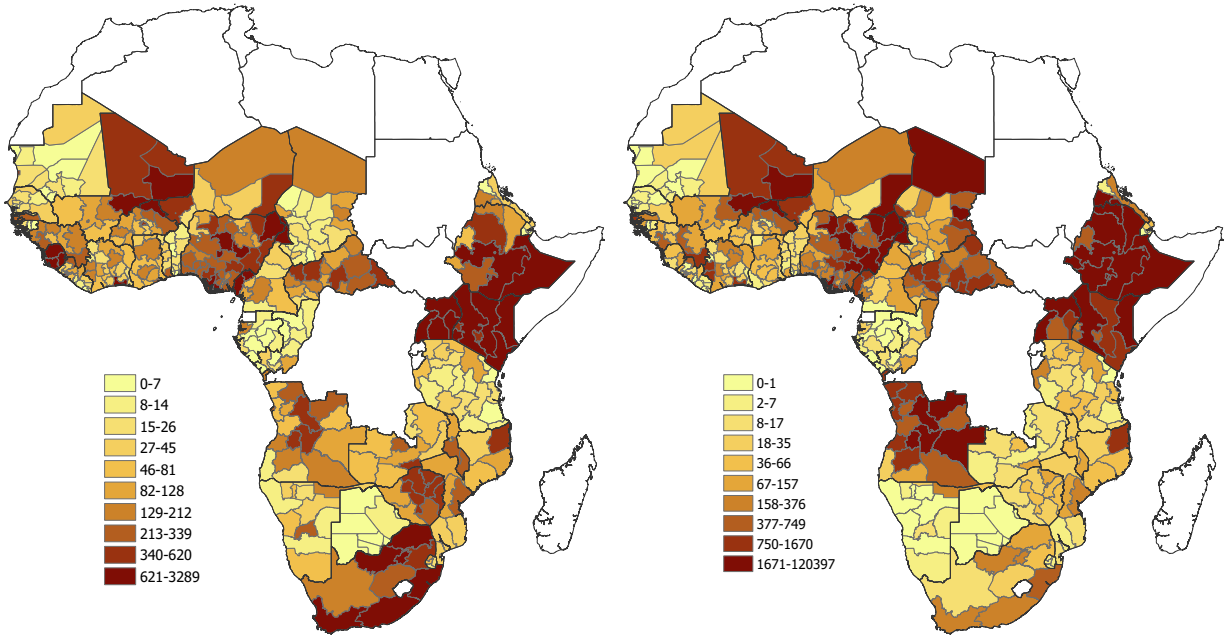


(a) BGI(1) (b) BGI(5)

(c) BGI(9) (d) BGI(13)

Figure 2: Spatial distribution of ethnic inequality in electricity access.

Notes. The legend in panel (d) applies to values of BGI(k) in all four panels.



(a) Events (b) Fatalities

Figure 3: Regional distribution of conflict events and fatalities from ACLED.

reports, and records from humanitarian agencies. The coverage of Sub-Saharan Africa starts in January 1997 and we use the data up to June 2020. ACLED events are geographically precise to the first subnational administrative level, which makes this source well-suited for our analysis. The dataset also reports the number of fatalities per conflict event. Figure 3 maps the counts of all ACLED events and fatalities for regions in our sample, reflecting the frequency and severity of conflict, respectively. There is a high within- and cross-country variation in both metrics, and they are mildly positively correlated.

In our robustness analysis, we use two additional data sources that focus on specific types of conflict, namely, the Uppsala Conflict Data Program (Sundberg and Melander, 2013), or UCDP, and the Social Conflict Analysis Database (Salehyan et al., 2012), or SCAD. In contrast to ACLED's broad coverage, UCDP only records deadly incidents associated with civil wars, and each armed conflict in the database resulted in at least 25 battle-related deaths in a calendar year. SCAD, on the other hand, excludes events that are linked to major conflicts and focuses on smaller-scale social unrest including demonstrations, riots, strikes, protests, and communal conflict. Summary statistics for all conflict variables are reported in appendix table D.1.

3 Linking regional ethnic diversity and conflict

3.1 Empirical framework

Following Michalopoulos and Papaioannou (2016), we estimate negative binomial models of conflict incidence and severity as functions of local ethnic diversity and other characteristics.[19] Our baseline estimating equation is

$$y_{ic} = \exp(\alpha_c + \beta D_{ic} + \mathbf{X'_{ic}}\boldsymbol{\gamma} + \varepsilon_{ic}),$$

where y_{ic} is one of conflict measures for region i of country c, D_{ic} is one of diversity indices from section 2.2, $\mathbf{X_{ic}}$ is a vector of regional control variables described below, α_c is the full set of country fixed effects, and ε_{ic} is the idiosyncratic error term. We estimate the model for 13 versions of each diversity index, corresponding to different levels of linguistic aggregation. This way we compare the roles of diversity based on relatively recent versus deep-rooted ethnolinguistic cleavages. The coefficient of interest is β and, given the presence of fixed effects, it is identified using within-country variation.

We control for a variety of relevant characteristics that may confound the main relationship of interest. Based on their importance for economic performance, conflict, and ethnolinguistic diversity (Michalopoulos, 2012), we account for both spatial average and standard deviation of agricultural suitability of land using the caloric suitability index of Galor and Özak (2016). We further include a number of geographic controls that have been shown to correlate with socioeconomic outcomes and/or population diversity, including terrain ruggedness (Nunn and Puga, 2012), malaria suitability index (Kiszewski et al., 2004; Cervellati et al., 2019), log of distance to the coast, log of surface area, and absolute latitude of region's centroid (Mitton, 2016). Given recent work on the role of genetic diversity in driving conflict (Arbatlı et al., 2020), we also add the log of distance from Addis Ababa to our list of baseline controls. To account for the potential importance of natural resource endowments (Berman et al., 2017), we include indicators for the presence of oil or gas fields and diamond mines. We also compute an indicator capturing the regional presence of ethnic homelands partitioned by state borders (Michalopoulos and Papaioannou, 2016). Additional, endogenous controls, including metrics of overall regional development and urbanization rate, are employed in robustness checks of section 3.4. Detailed definitions of all variables are provided in appendix A.

[19] The estimates from Poisson regressions are qualitatively similar. See appendix E.2 for a subset of our key results reproduced in this alternative framework.

3.2 Baseline results

For convenience and visual clarity, we present our results in a compact graphical form similar to Gershman and Rivera (2018). Specifically, for each diversity indicator, we build a diagram showing the estimates of interest at odd-numbered tiers of linguistic aggregation, from 1 to 11 for both conflict outcomes (events and fatalities).[20] For ease of interpretation, we standardize diversity indices to have zero mean and unit standard deviation prior to estimation. We further transform the estimated β coefficients into incidence rate ratios (IRR), which are reported in the diagrams along with respective 95% confidence intervals based on robust standard errors.

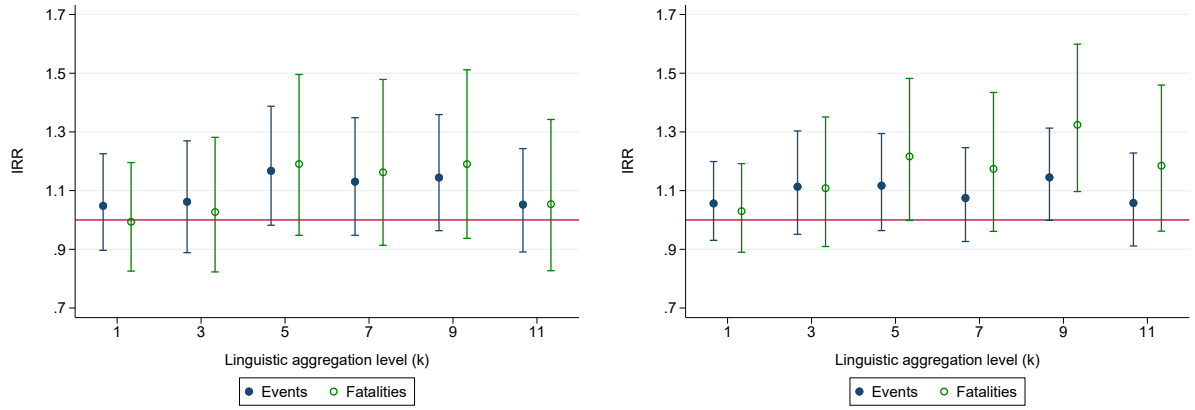
Figure 4 presents estimation results for four indices of ethnic inequality introduced earlier. The vast majority of point estimates are positive, and they tend to be somewhat larger in magnitude for conflict fatalities. However, almost none of these estimates are statistically significant at the 5% level, as can be seen from the overlap of respective confidence intervals and the reference horizontal line at IRR equal to 1 (marking the absence of association). The results for BGI in access to electricity suggest that, other things equal, a one-standard-deviation increase in BGI(9) is associated, on average, with a 15% rise in the number of conflict events and more than a 30% increase in fatalities.

Our results for indices of ethnic inequality based on other indicators, presented in appendix E.1, are largely similar. Just as in the case of car and television ownership, there is no robust significant relationship for BGI in access to other "luxury" goods such as flush toilet and finished flooring, as well as high-quality water sources. Interestingly, ethnic inequality in radio ownership appears to be most strongly related to conflict, at least for some linguistic aggregation levels. However, as shown below, even those few BGI indices that are significant in these baseline regressions lose their "horse races" against ELF.

Figure 5 shows our results for ethnolinguistic fractionalization and polarization. The estimates for ELF are positive, highly statistically significant, and large, especially for $k \in [3, 9]$. At these aggregation levels, a one-standard-deviation greater ELF is associated with about 25-35% higher incidence of conflict events and 30-50% more fatalities. These results suggest that deep-rooted diversity matters more, although the estimates become smaller at tier 1, perhaps reflecting the lack of sufficient variation in ELF at that level.

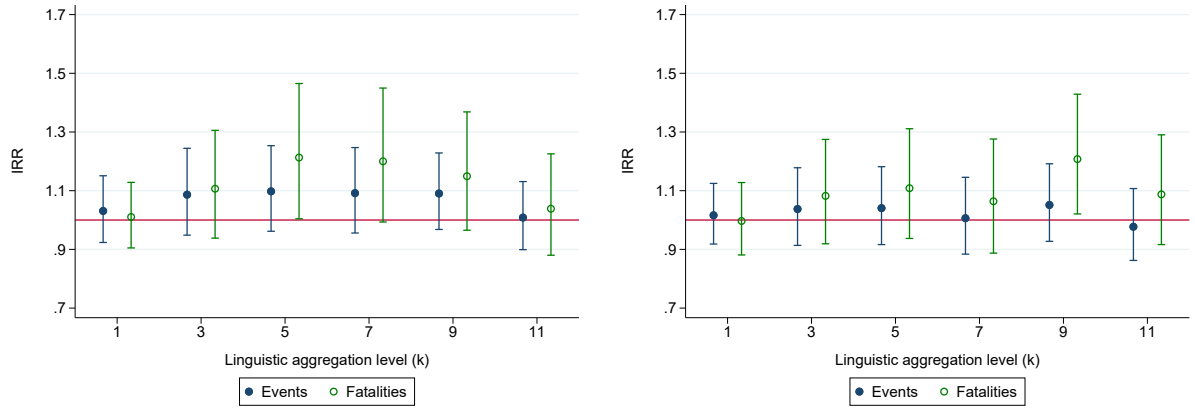
The relationship between ELP and conflict is quite different for $k \geqslant 7$, with point estimates mostly insignificant. For higher levels of linguistic aggregation, ELP and ELF

[20] We omit the results for even-numbered tiers to make the figures more transparent. The reported estimates are fully representative of the key patterns. The results for tiers 11 and 13 are virtually identical, and we omit the latter.



(a) Years of education

(b) Access to electricity



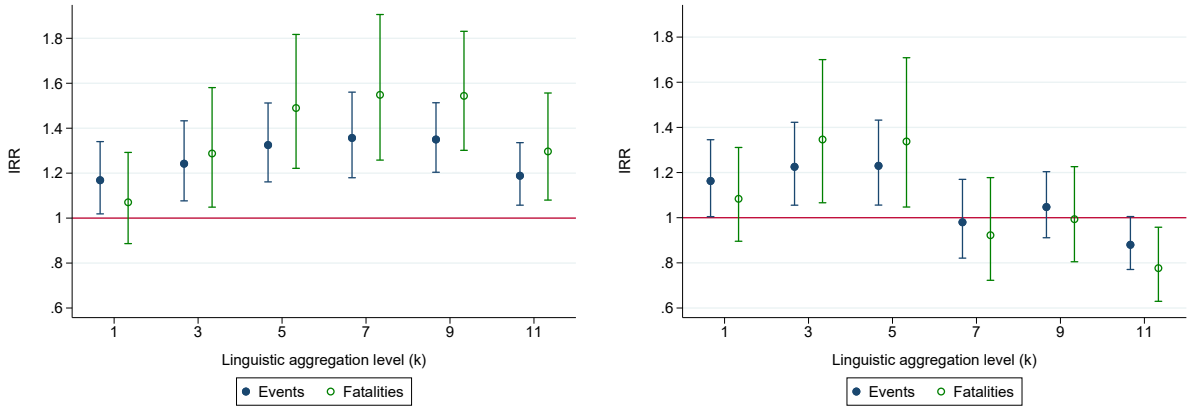
(c) Television ownership

(d) Car ownership

Figure 4: Subnational ethnic inequality and conflict.

Notes. Each panel presents incidence rate ratios, along with 95% confidence intervals, based on robust standard errors. For each of the six reported levels of linguistic aggregation, a negative binomial regression is estimated, where the outcome is the sum of either conflict events or fatalities, and the right-hand-side variable of interest is a BGI(k) index capturing ethnic inequality in the dimension indicated in the figure subtitle. All regressions include country fixed effects and baseline controls described in the main text. The number of observations is 391.

results are similar, due to a well-known tendency of these indices to be very highly correlated under restricted value range, which is what happens for smaller k (Gershman and Rivera, 2018). In subsequent tests, we generally found that the estimates for ELP at larger k values are not robust and we limit the remaining exposition to the comparative analysis of ELF and BGI in their relation to conflict.



(a) ELF

(b) ELP

Figure 5: Subnational fractionalization, polarization, and conflict.

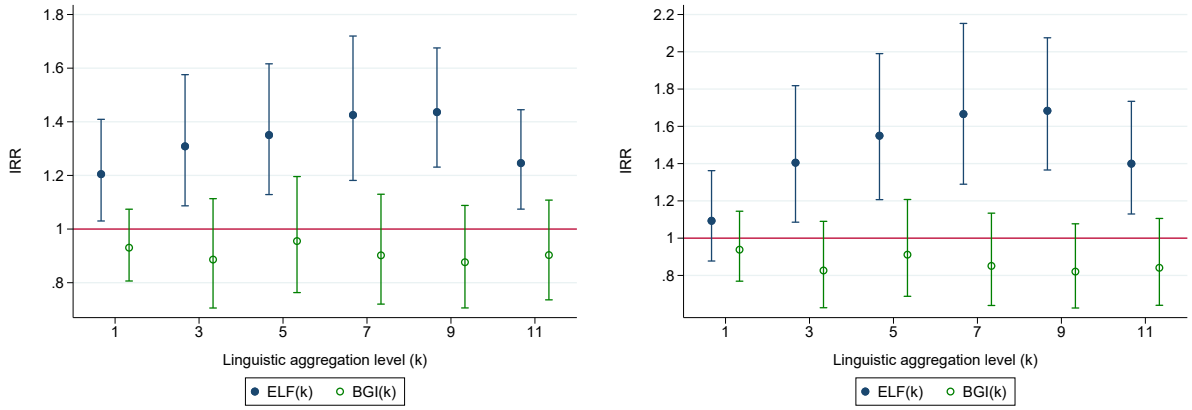
Notes. Each panel presents incidence rate ratios, along with 95% confidence intervals, based on robust standard errors. For each of the six reported levels of linguistic aggregation, a negative binomial regression is estimated, where the outcome is the sum of either conflict events or fatalities, and the right-hand-side variable of interest is ELF(k) in panel (a) and ELP(k) in panel (b). All regressions include country fixed effects and baseline controls described in the main text. The number of observations is 391.

Among the control variables included in our baseline specification, several turned out to be statistically significant throughout the analysis, with coefficient signs largely consistent with earlier studies of conflict in Sub-Saharan Africa. Absolute latitude and the indicator of ethnic partitioning are positively associated with conflict, although the latter is somewhat less robust to model specification in conflict events regressions. On the other hand, distances to Addis Ababa and coastline are negatively related to conflict incidence.

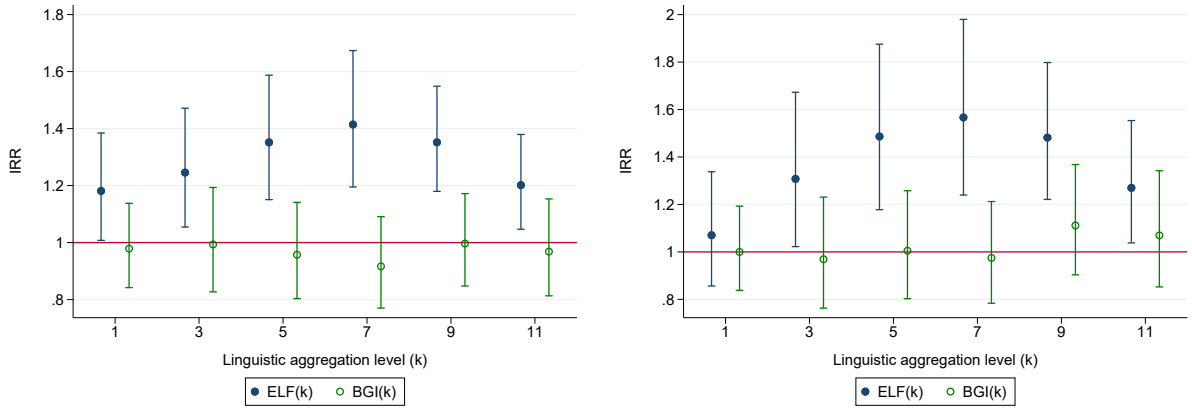
3.3 ELF vs. BGI

Given our baseline results, we attempt to disentangle the relationship between subnational diversity and conflict focusing on ELF(k) and BGI(k). This is especially important in light of the recent debate regarding the relative importance of these two metrics in predicting socioeconomic outcomes and the substantial positive correlation between them reported in section 2.2. Recall that, while ELF captures the ethnolinguistic fragmentation of population, BGI measures socioeconomic differences between groups.

We conduct this comparison by estimating "horse race" regressions. Specifically, we estimate the same negative binomial model as earlier, but simultaneously include both ELF and BGI indices, measured at the same level of linguistic aggregation. We present the



(a) BGI in years of education and ELF as predictors of conflict events (left) and fatalities (right)



(b) BGI in electricity access and ELF as predictors of conflict events (left) and fatalities (right)

Figure 6: "Horse race" regressions for ELF and BGI.

Notes. Each panel presents incidence rate ratios, along with 95% confidence intervals, based on robust standard errors. For each of the six reported levels of linguistic aggregation, a negative binomial regression is estimated, where the outcome is the sum of conflict events or fatalities. The right-hand side includes both ELF(k) and BGI(k) for years of education (panel a) electricity access (panel b), along with country fixed effects and baseline controls described in the main text. The number of observations is 391.

results for BGI in years of education and electricity access, and relegate similar diagrams for other BGI indices to appendix E.3.[21]

Figure 6 presents the incidence rate ratios for ELF and BGI from the "horse race" regressions. The clear "winner" here is ELF, with coefficient estimates becoming even stronger relative to the baseline in panel (a) of figure 5. In contrast, estimates for BGI are mostly negative and/or close to zero, and statistically insignificant across the board. In

[21] These supplementary diagrams focus on conflict events as an outcome variable. The results for fatalities are qualitatively similar.

other words, ethnic inequality in electricity access and other outcomes (figure E.3) appears to be unrelated to conflict incidence or severity after accounting for ELF. The IRRs for ELF are generally larger for $k < 11$ implying that deeper ethnolinguistic cleavages matter more for generating conflict. As mentioned earlier, weaker results for $k = 1$ could simply reflect the lack of substantial variation in diversity when aggregating ethnic groups to the level of major language families.[22]

In sum, our analysis suggests that ethnolinguistic fragmentation, particularly based on deeper linguistic cleavages, has a much stronger positive association with conflict than ethnic inequality. Identity-based distinctions, rather than the presence of larger group-level differences in socioeconomic status, appear to be a better predictor of conflict.

3.4 Robustness tests

We next show the robustness of our key results. Although our baseline specification controls for a host of relevant regional characteristics and country fixed effects, there may still be other important omitted variables. Tables 1 and 2 show that our qualitative results hold in model specifications accounting for additional factors, whether ELF and BGI enter the regression equation separately or jointly.

The first column of table 1 displays our baseline estimates from conflict events regressions, previously reported in graphical form in figures 4b and 5a. The following three columns augment these specifications by including, one by one, three metrics of average level of development and/or local public goods provision, namely, average literacy rate, electricity access, and international wealth index (IWI) of Smits and Steendijk (2015). This is done to confirm that our results are not driven by the negative relationship between deep-rooted diversity and average access to local public goods (Gershman and Rivera, 2018). The estimates show that coefficients on both ELF and BGI in electricity access are largely unaffected by the literacy and electrification variables. Controlling for IWI in column 4 has a more substantial impact on ELF estimates (without altering the qualitative results), although, due to data availability constraints, the sample size is also reduced by 15 regions.

In column 5, we add the rate of urbanization to our set of baseline controls. Urban areas tend to be regions of highest ethnic diversity and also hubs of development and local public goods provisions. Our estimates for ELF become smaller and ELF(11) loses its statistical significance, but the qualitative pattern remains intact: deep-rooted diversity is positively associated with conflict incidence. Interestingly, controlling for urbanization rate seems to mildly strengthen the BGI estimates.

[22] As shown in appendix E.2, while Poisson regression estimates are very similar for conflict events, they are substantially larger in the case of fatalities, especially for $k = 1$.

Table 1: Robustness to alternative specifications, separately for ELF and BGI

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ELF(1)	1.169**	1.221***	1.189**	1.123	1.140*	1.173**	1.167**
	(0.026)	(0.005)	(0.017)	(0.113)	(0.074)	(0.024)	(0.028)
ELF(3)	1.242***	1.292***	1.261***	1.216***	1.217***	1.250***	1.237***
	(0.003)	(0.000)	(0.002)	(0.009)	(0.008)	(0.003)	(0.003)
ELF(5)	1.325***	1.361***	1.330***	1.278***	1.254***	1.327***	1.323***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
ELF(7)	1.357***	1.377***	1.352***	1.288***	1.258***	1.349***	1.355***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)	(0.000)	(0.000)
ELF(9)	1.350***	1.365***	1.340***	1.296***	1.256***	1.341***	1.350***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ELF(11)	1.188***	1.191***	1.168***	1.122*	1.093	1.173***	1.183***
	(0.004)	(0.003)	(0.010)	(0.058)	(0.151)	(0.007)	(0.006)
BGI(1)	1.056	1.079	1.072	1.071	1.070	1.081	1.060
	(0.396)	(0.255)	(0.287)	(0.302)	(0.318)	(0.242)	(0.374)
BGI(3)	1.113	1.154*	1.135	1.133	1.147	1.150*	1.112
	(0.181)	(0.086)	(0.121)	(0.134)	(0.101)	(0.089)	(0.190)
BGI(5)	1.117	1.160*	1.154*	1.149*	1.143*	1.177**	1.111
	(0.141)	(0.057)	(0.064)	(0.075)	(0.087)	(0.039)	(0.160)
BGI(7)	1.075	1.130	1.119	1.078	1.110	1.151*	1.068
	(0.341)	(0.124)	(0.141)	(0.333)	(0.165)	(0.083)	(0.383)
BGI(9)	1.145*	1.207**	1.205***	1.181**	1.195***	1.244***	1.141*
	(0.052)	(0.010)	(0.008)	(0.018)	(0.009)	(0.004)	(0.060)
BGI(11)	1.058	1.109	1.118	1.094	1.106	1.146*	1.051
	(0.460)	(0.194)	(0.150)	(0.242)	(0.176)	(0.099)	(0.520)
Observations	391	391	391	376	391	391	391
Literacy		✓					
Electrification			✓				
IWI				✓			
Urbanization					✓		
Gini coefficient						✓	
Religious frac.							✓

Notes. a) The top panel of the table shows estimates from 43 separate negative binominal regressions of ACLED conflict events on ELF(k). The bottom panel does the same for BGI(k) in access to electricity. All regressions contain baseline controls described in the main text. IRRs are reported for each regression, along with p-values (based on robust standard errors) in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% level, respectively. b) Additional controls listed at the bottom of the table are regional averages for literacy rate, electricity access, international wealth index (IWI), urbanization rate, Gini coefficient in electricity access, and an index of religious fractionalization.

Table 2: Robustness to alternative specifications: ELF vs. BGI horse races

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ELF(1)	1.181**	1.227**	1.196**	1.111	1.133	1.172*	1.177**
	(0.040)	(0.012)	(0.033)	(0.205)	(0.134)	(0.054)	(0.045)
BGI(1)	0.979	0.990	0.989	1.022	1.014	1.002	0.983
	(0.779)	(0.897)	(0.883)	(0.769)	(0.856)	(0.975)	(0.828)
ELF(3)	1.246***	1.281***	1.256***	1.199**	1.193**	1.235**	1.240**
	(0.010)	(0.003)	(0.008)	(0.035)	(0.038)	(0.014)	(0.011)
BGI(3)	0.993	1.022	1.010	1.035	1.049	1.027	0.995
	(0.943)	(0.822)	(0.917)	(0.719)	(0.613)	(0.779)	(0.955)
ELF(5)	1.351***	1.366***	1.334***	1.269***	1.241***	1.322***	1.350***
	(0.000)	(0.000)	(0.000)	(0.004)	(0.008)	(0.001)	(0.000)
BGI(5)	0.957	0.991	0.994	1.017	1.025	1.008	0.958
	(0.627)	(0.919)	(0.943)	(0.856)	(0.786)	(0.928)	(0.628)
ELF(7)	1.414***	1.404***	1.377***	1.318***	1.260***	1.367***	1.413***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.009)	(0.000)	(0.000)
BGI(7)	0.917	0.959	0.961	0.951	0.997	0.974	0.917
	(0.327)	(0.650)	(0.658)	(0.576)	(0.973)	(0.784)	(0.328)
ELF(9)	1.352***	1.341***	1.311***	1.267***	1.209***	1.303***	1.352***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.008)	(0.000)	(0.000)
BGI(9)	0.997	1.046	1.054	1.058	1.092	1.072	0.997
	(0.967)	(0.601)	(0.528)	(0.500)	(0.271)	(0.456)	(0.967)
ELF(11)	1.202***	1.186**	1.153**	1.107	1.066	1.151*	1.197**
	(0.009)	(0.015)	(0.049)	(0.167)	(0.380)	(0.060)	(0.012)
BGI(11)	0.968	1.014	1.037	1.040	1.071	1.050	0.968
	(0.719)	(0.883)	(0.698)	(0.670)	(0.425)	(0.631)	(0.713)
Observations	391	391	391	376	391	391	391
Literacy		✓					
Electrification			✓				
IWI				✓			
Urbanization					✓		
Gini coefficient						✓	
Religious frac.							✓

Notes. a) The table shows estimates from 43 separate negative binominal horse race regressions of ACLED conflict events on ELF(k) and BGI(k). Each of the six panels displays results for a different level of linguistic aggregation k. All regressions contain baseline controls described in the main text. IRRs are reported for each regression, along with p-values (based on robust standard errors) in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% level, respectively. b) Additional controls listed at the bottom of the table are regional averages for literacy rate, electricity access, international wealth index (IWI), urbanization rate, Gini coefficient in electricity access, and an index of religious fractionalization.

Finally, the last two columns report estimates after including regional measures of overall (rather than between-group) inequality in electricity access and an index of religious fractionalization from Gershman and Rivera (2018). This has very little effect on the ELF estimates, but some BGI estimates appear stronger after including the "raw" Gini coefficient. All additional control variables, with the exception of religious fractionalization, are statistically significant when included separately, and only Gini coefficient is negatively related to conflict events. However, when all variables are accounted for simultaneously, only urbanization rate retains a statistically significant (positive) association with conflict.

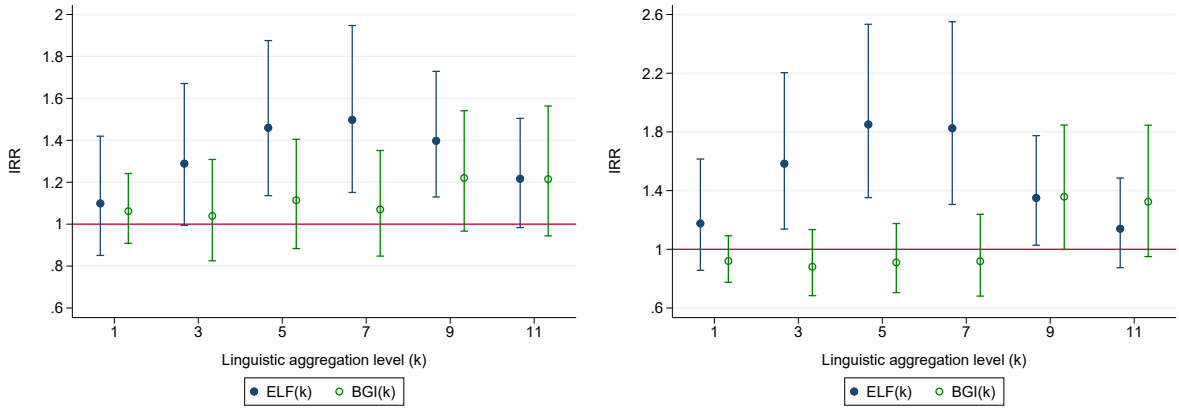
Table 2 shows that our horse race results are likewise robust to alternative model specifications. The first column shows the baseline estimates corresponding to panel (b) of figure 6, while the remaining columns reflect the same sequence of sensitivity tests as in table 1. The dominance of ELF over BGI as a predictor of conflict is clear: the coefficients on ELF(3)–ELF(9) are sizable and statistically significant, whereas BGI is insignificant in every single regression.

Although we cannot completely rule out the presence of some unobservable factors that may still affect our results, their robustness to a variety of potentially confounding characteristics gives us confidence. An additional empirical challenge to the observed relationship between ELF and conflict is population sorting. Indeed, it is plausible that individuals seeking security would try to leave conflict-ridden regions and avoid moving there. This would result in a *negative* association between diversity and conflict, in which case our estimates understate the true positive effect of ELF.

Furthermore, as documented in Gershman and Rivera (2018), subnational ELF in Sub-Saharan Africa is extremely persistent over time, suggesting negligible effects of population sorting. Specifically, for several countries with reliable data, they show that the correlations between ELF(k) values over several decades are well above 0.9. This persistence is stronger for lower values of k and appears just as strong for smaller administrative subdivisions.

Our results are also robust to alternative measures of conflict incidence and intensity, as shown in figure 7. Panels (a) and (b) illustrate the estimates from "horse race" regressions for ELF and BGI (in electricity access) when UCDP data on deadly conflicts are used to construct the dependent variables. The overall pattern is roughly similar to that observed in panel (b) of figure 6 for ACLED data, but implies a stronger magnitude of association between deep-rooted ELF and conflict.[23] In panels (c) and (d), we report analogous re-

[23]In an additional robustness check, we use the Ethnic Power Relations (EPR) database (Vogt et al., 2015) to extract only those conflict events that have an ethnic dimension. In such cases, at least one side of conflict claimed to represent, recruited from, or was supported by a specific ethnic group. Our results for this subset of conflict events and fatalities are qualitatively similar to those for the unrestricted UCDP sample presented in figure 7.



(a) UCDP Events

(b) UCDP Fatalities

(c) SCAD Events

(d) SCAD Fatalities

Figure 7: ELF vs. BGI "horse race" regressions for alternative conflict data.

Notes. Each panel presents incidence rate ratios, along with 95% confidence intervals, based on robust standard errors. For each of the six reported levels of linguistic aggregation, a negative binomial regression is estimated, where the outcome is the sum of conflict events or fatalities based on UCDP data (panels a and b) or SCAD data (panels c and d). The right-hand side includes both ELF(k) and BGI(k) for electricity access, along with country fixed effects, baseline controls, and urbanization rate. The number of observations is 391.

sults for SCAD data covering a broader scope of social unrest incidents, both violent and non-violent. Once again, the qualitative pattern is similar to our baseline, with larger IRR estimates. Thus, fragmentation (particularly when measured for deeper linguistic cleavages) but not ethnic inequality is strongly positively associated with a broad spectrum of conflict measures.[24]

[24]The regression results are qualitatively the same when using BGI in other socioeconomic outcomes.

4 Concluding remarks

In this paper, we present a newly constructed dataset on subnational ethnolinguistic divisions in Sub-Saharan Africa, with a focus on ethnic inequality, and use it to revisit the relationship between diversity and conflict. We leverage microdata from large-scale household surveys that allow us to distinguish between various dimensions of ethnic inequality and accurately capture the ethnolinguistic composition of each region in our sample. Furthermore, we account for the depth of ethnolinguistic cleavages by integrating our data into the global linguistic tree model and computing diversity metrics at different levels of linguistic aggregation.

Our empirical analysis reveals a robust positive relationship between ethnolinguistic fractionalization and conflict, which is stronger for ELF measures based on deep-rooted linguistic cleavages. This underscores the importance of accounting for relatedness between groups when evaluating the role of diversity in societies.

In contrast to some recent research emphasizing the importance of ethnic-group-level differences in socioeconomic characteristics for driving conflict, we do not find systematic evidence to support this view at the subnational level. The relationship between our numerous measures of ethnic inequality and conflict intensity of severity is weak and lacks statistical significance, especially after accounting for ethnolinguistic fractionalization. Thus, fragmentation rather than inequality appears to be the more relevant aspect of regional ethnic divisions in relation to conflict.

Our findings have important implications for policies related to territorial administration, redistribution, and peace. The imposition of state borders during the colonial era had a lasting impact on African development, and the redrawing of internal administrative boundaries remains a contentious issue today. Policymakers facing this task should consider potential costs of social unrest associated with greater ethnolinguistic fractionalization in the resulting regions, with special attention to ancestral relationships between relevant groups. Similar considerations can inform the design of refugee resettlement programs and policies catering to migrant populations. The lack of association between conflict and ethnic inequality also casts doubt on the effectiveness of targeted local-level redistribution in reducing conflict.

Appendices

A Description of variables

Ethnic inequality and other diversity measures

The BGI, ELF, and ELP indices are described in detail in section 2.2. The variables underlying the BGI indices were defined as follows.

Education in years. Educational attainment in years was typically recorded based on the last completed year of schooling. If this direct measure is missing, years of education were calculated from the reported level of education (e.g., primary) using country-specific schooling requirements that were in place at the time of the survey.

Access to a high-quality water source. Indicator of household access to one of the following sources: water piped into dwelling or plot, water supplied via a water tanker, bottled water. Low-quality sources include private or community wells and taps, nearby natural sources.

Flush toilet. Indicator of household access to flush toilet (as opposed to pit latrines and other lower-quality options).

Finished floor. High-quality flooring indicator, equal to one if the floor is finished (e.g., with parquet, carpet, or tiles).

Electricity access. Indicator of household access to electricity.

Asset ownership. Indicators of household ownership of radio, television, and car.

Conflict variables

ACLED events and fatalities. The regional count of conflict events and fatalities from the Armed Conflict Location and Event Database (from January 1997 to June 2020). *Source:* `https://acleddata.com/data-export-tool/` and own calculations.

UCDP events and fatalities. The regional count of conflict events and fatalities from the Uppsala Conflict Data Program's Georeferenced Event Dataset (1997–2019). *Source*: `https://ucdp.uu.se/downloads/` and own calculations.

SCAD events and fatalities. The regional count of conflict events and fatalities from the Social Conflict Analysis Database (1997–2017). *Source*: `https://www.strausscenter.org/ccaps-research-areas/social-conflict/database/` and own calculations.

Control variables

Terrain ruggedness index. Index of terrain ruggedness as constructed by Nunn and Puga (2012) at 0.5-degree resolution, averaged across cells in each region. *Source*: Gershman and Rivera (2018).

Caloric suitability index. An index capturing potential agricultural output (measured in calories) based on crops that were available for cultivation in the post-1500CE era and available at the 5 arc-minute resolution (Galor and Özak, 2016). We calculate the mean and standard deviation of the index based on cell values within each region. *Source:* `https://ozak.github.io/Caloric-Suitability-Index` and own calculations.

Distance to the coastline. Great circle distance from a region's centroid to the closest location on the coastline. *Source:* `http://www.naturalearthdata.com` and own calculations.

Distance to Addis Ababa: Great circle distance from a region's centroid to Addis Ababa. *Source*: own calculations

Oil and gas field indicator. A dummy variable indicating the presence of an oil or gas field in the region. *Source*: PETRODATA dataset (version 1.2) and own calculations.

Diamond occurence. A dummy variable indicating the occurence of diamonds in the region. *Source*: DIADATA dataset and own calculations.

Partitioned ethnic homeland indicator. A dummy variable equal to one, if a region intersects an ethnic homeland partitioned by state borders, as defined in Michalopoulos and Papaioannou (2016). *Source*: own calculations.

Malaria stability index. An index measuring the stability of malaria transmission based on regionally dominant vector mosquitoes, available at 0.5-degree resolution. We calculate the mean across cells in each region. *Source*: Kiszewski et al. (2004) and own calculations.

Literacy rate. Share of region's adult population (aged 15–49 years) that is literate. A person is considered literate if she can read at least part of a standard sentence or has attended secondary school. *Source*: Gershman and Rivera (2018).

Electrification. Share of region's households that have access to electricity. *Source*: Gershman and Rivera (2018).

International wealth index (IWI). Wealth index, as proposed by Smits and Steendijk (2015), averaged across households within relevant regions. *Source*: Gershman and Rivera (2018).

Urbanization rate. Share of region's households that live in urban areas. *Source*: Gershman and Rivera (2018).

Religious fractionalization. Regional index of fractionalization based on a four-way classification for religious affiliation: Christianity, Islam, traditional religion, and none. *Source*: Gershman and Rivera (2018).

Gini coefficient for electricity access. Gini coefficient measuring inequality in household access to electricity at the regional level. *Source*: own calculations.

B Microdata used to measure population diversity

Table B.1: Countries, data sources, ethnolinguistic groups, and subnational regions

Country	Primary source	Secondary source	Groups (P/S)	Regions
Angola	DHS (2016)		10	18
Benin	IPUMS (2013)	DHS (2012)	42/8	12
Botswana	IPUMS (2011)		12	9
Burkina Faso	IPUMS (2006)	MICS (2006)	27/18	13
Cameroon	DHS (2018)		138	10
Central African Rep.	MICS (2010)		9	17
Chad	DHS (2015)		19	20
Rep. of the Congo	DHS (2012)		11	12
Côte d'Ivoire	MICS (2006)		50	19
Djibouti	MICS (2006)		3	6
Eritrea	DHS (2002)		10	6
Eswatini	DHS (2007)		1	4
Ethiopia	IPUMS (2007)	DHS (2011)	63/48	11
Gabon	DHS (2012)		8	9
Gambia	DHS (2013)		8	8
Ghana	IPUMS (2010)	DHS (2008)	38/8	10
Guinea	DHS (2012)		6	8
Guinea-Bissau	MICS (2018)		8	9
Kenya	DHS (2014)		20	8
Liberia	IPUMS (2008)	DHS (2013)	17/17	15
Malawi	IPUMS (2008)	DHS (2010)	9/7	3
Mali	IPUMS (2009)	MICS (2010)	14/16	9
Mauritania	MICS (2007)		4	13
Mozambique	MICS (2008)		22	11
Namibia	DHS (2013)		8	13
Niger	DHS (2006)		8	8
Nigeria	DHS (2013)		192	37
Senegal	IPUMS (2013)	DHS (2011)	19/6	11
Sierra Leone	IPUMS (2004)	MICS (2017)	15/14	4
South Africa	IPUMS (2011)	DHS (2016)	11/11	9
Tanzania	DHS (1992)		98	21
Togo	IPUMS(2010)	DHS (2014)	31/5	5
Uganda	IPUMS (2014)	DHS (2016)	41/37	4
Zambia	IPUMS (2010)	DHS (2014)	32/30	9
Zimbabwe	DHS (2011)		2	10

Notes. a) DHS is Demographic and Health Surveys, MICS is Multiple Indicator Cluster Surveys, IPUMS is Integrated Public Use Microdata Series (subsamples of national censuses). b) Column 4 lists the number of unique ethnolinguistic groups in our primary (P) and secondary (S) data sources. d) See Gershman and Rivera (2018) for details on the definitions of subnational boundaries.

C Computing BGI(k)

As explained in the main text, we use *Ethnologue*'s language tree model to measure ethnic inequality at different levels of linguistic aggregation. Here, we illustrate the process using the Gash-Barka region of Eritrea as example. In our sample, there are 7 ethnic groups in this region that uniquely match to 7 languages in *Ethnologue*. These languages (in boxes) and the hypothesized ancestral relationships between them are illustrated in figure C.1, which is the relevant linguistic subtree extracted from *Ethnologue*. As can be seen, all languages spoken in the Gash-Barka region belong to two major families, Afro-Asiatic and Nilo-Saharan. This is the deepest (tier 1) cleavage in *Ethnologue*'s model beyond which it only identifies the "Proto-Human" root referring to the hypothetical common ancestor of all languages. Beyond tier 1, the tree branches out further ultimately giving rise to contemporary languages. The longer the path shared by two languages before they diverge, the more closely related they are. In our case, Tigre and Tigrigna are the closest two languages as they share 5 branches starting from tier 1, whereas any pair of languages representing two distinct major language families are the least related since they share no common branches.

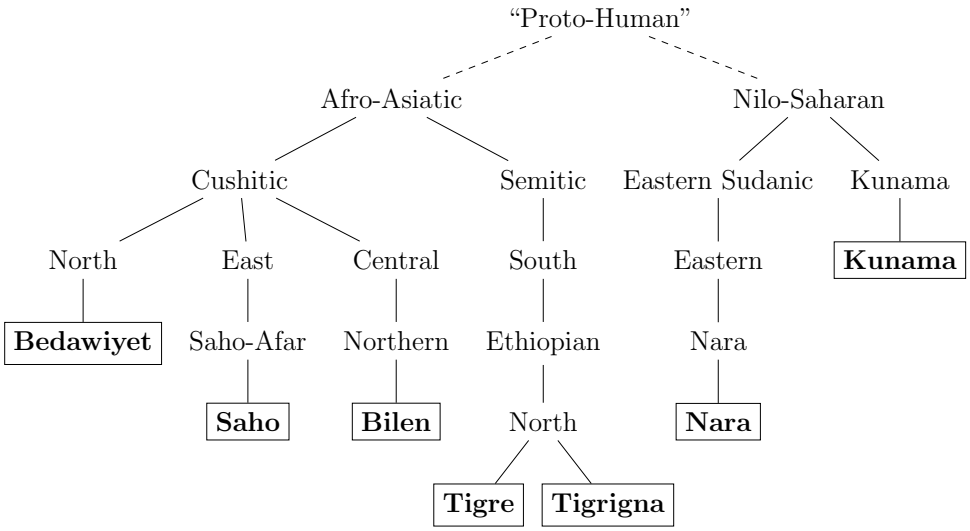


Figure C.1: Ethnolinguistic tree for languages spoken in the Gash-Barka region of Eritrea

To account for linguistic relatedness we follow the aggregation method from Desmet et al. (2012). The first step in this approach is to create "artificial" ancestral groups to equate the number of branches from the root to each extant language. The outcome of this process is shown in figure C.2, where the added ancestral groups are marked with asterisks. With each of the 6 possible tiers of the tree now properly defined, we can aggregated languages

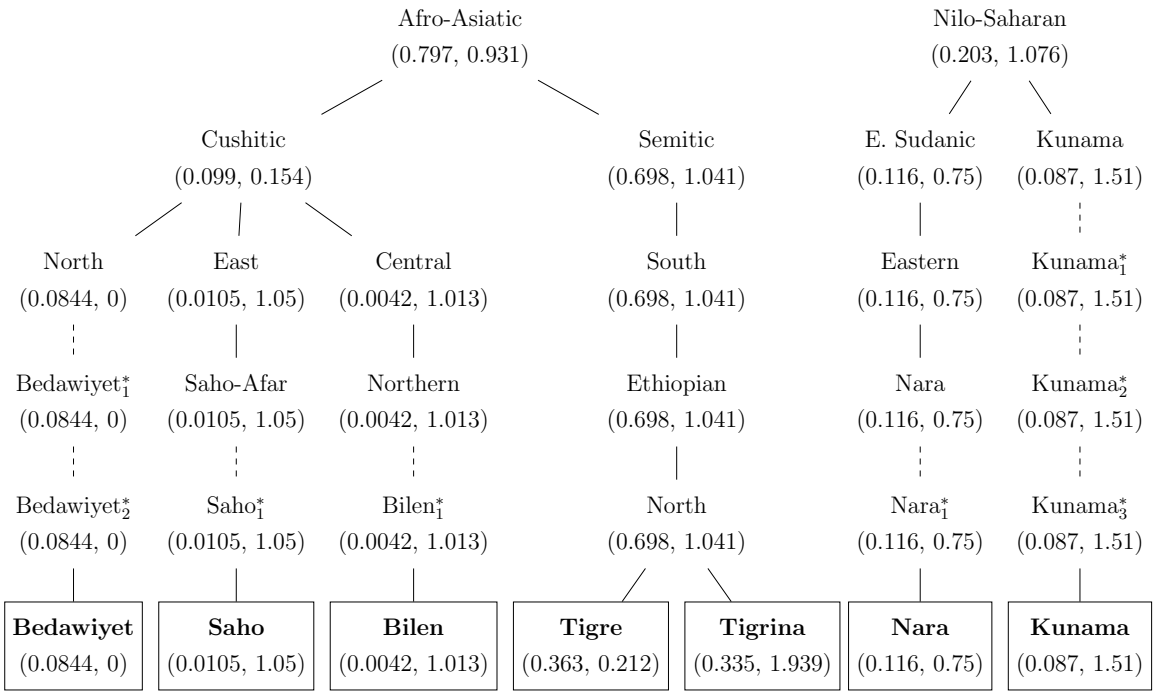


Figure C.2: Extended ethnolinguistic tree for the Gash-Barka region of Eritrea

Notes. Population shares and mean educational attainment reported in parentheses.

to the level of ancestral groups. For example, at tier 5, Tigre and Tigrigna "merge" into a single North subgroup of Ethiopian languages and at tier 2, Bedawiyet, Saho, and Bilen all "merge" into a single Cushitic subfamily. Thus, by construction, there are fewer distinct groups at higher levels of linguistic aggregation.

Figure C.2 reports other crucial elements required to measure ethnic inequality, namely, the (rounded) regional population shares and mean levels of educational attainment for each (aggregated) group. Note that both depend on the tier of the tree. For example, at tier 5, where Tigre and Tigrigna merge into North, this group becomes the largest, constituting about 70% of the region's population, and its mean level of educational attainment is the weighted average of the respective values for the two constituent groups. We compute population shares and mean outcomes for all relevant ancestral groups, as shown in the figure. The overall regional mean educational attainment level in Gash-Barka is 0.96 years. Next, we simply use the BGI formula at each tier of the tree. For example, at tier 1, we have $\text{BGI}(1) = (2\times0.797\times0.203\times|0.931-1.076|)/(2\times0.96) \approx 0.024$. Similarly, $\text{BGI}(2) \approx 0.145$, $\text{BGI}(3) = \text{BGI}(4) = \text{BGI}(5) \approx 0.148$, and $\text{BGI}(6) \approx 0.449$.

As mentioned in the main text, in our full sample, there are 13 tiers of aggregation since for some languages we count up to 12 steps from the root to the bottom of the tree. In the case of Gash-Barka, all indices for tiers 6 through 13 will be identical: additional tiers in the extended tree will simply be "copies" of tier 6.

C.1 BGI(k) is non-decreasing in k

Consider tier k of a linguistic tree. Assume that, at tier k, there are three distinct ethnolinguistic groups with population shares s_1, s_2, and s_3 and average outcomes $\bar{w}_1$, $\bar{w}_2$, and $\bar{w}_3$. Assume further that at tier $k-1$, groups 2 and 3 merge into a single ancestral group. Its population share is $s_2 + s_3$ and its average outcome is $\bar{w}_{23}$, the weighted average of $\bar{w}_2$ and $\bar{w}_3$:

$$\bar{w}_{23} = \frac{s_2}{s_2 + s_3} \cdot \bar{w}_2 + \frac{s_3}{s_2 + s_3} \cdot \bar{w}_3.$$

We claim that $\text{BGI}(k) \geqslant \text{BGI}(k-1)$ in the same region. In our example,

$$\text{BGI}(k) = \frac{1}{2\bar{w}} \cdot (s_1 s_2 |\bar{w}_1 - \bar{w}_2| + s_1 s_3 |\bar{w}_1 - \bar{w}_3| + s_2 s_3 |\bar{w}_2 - \bar{w}_3|),$$
$$\text{BGI}(k-1) = \frac{1}{2\bar{w}} \cdot (s_1 (s_2 + s_3) |\bar{w}_1 - \bar{w}_{23}|).$$

Plug in the earlier expression for w_{23} and rearrange to get

$$s_1(s_2 + s_3)|\bar{w}_1 - \bar{w}_{23}| = s_1 |s_2 \bar{w}_1 + s_3 \bar{w}_1 - s_2 \bar{w}_2 - s_3 \bar{w}_3| = |s_1 s_2 (\bar{w}_1 - \bar{w}_2) + s_1 s_3 (\bar{w}_1 - \bar{w}_3)|$$

By triangle inequality,

$$\text{BGI}(k-1) = \frac{1}{2\bar{w}} \cdot |s_1 s_2 (\bar{w}_1 - \bar{w}_2) + s_1 s_3 (\bar{w}_1 - \bar{w}_3)| \leqslant \frac{1}{2\bar{w}} \cdot (s_1 s_2 |\bar{w}_1 - \bar{w}_2| + s_1 s_3 |\bar{w}_1 - \bar{w}_3|) \leqslant \text{BGI}(k).$$

D Summary statistics

Table D.1: Summary statistics for conflict measures

Variable	Mean	St. dev.	Min	25th pct.	Median	75th pct.	Max
ACLED events	242.5	459.8	0	20	79	262	3289
ACLED fatalities	1059.8	6786.6	0	9	57	489	120397
UCDP events	38.2	137.0	0	0	2	23	2074
UCDP fatalities	427.9	1630.9	0	0	9	242	23453
SCAD events	12.9	30.7	0	1	3	10	262
SCAD fatalities	70.0	276.0	0	0	2	26	4234

Notes. ACLED refers to the Armed Conflict and Location Event Database. UCDP stands for the Uppsala Conflict Data Program. SCAD refers to the Social Conflict Analysis Database.

Table D.2: Summary statistics for population diversity indices

Tier k	Education BGI(k)	Electricity BGI(k)	Television BGI(k)	Car BGI(k)	Water BGI(k)	Toilet BGI(k)	Floor BGI(k)	Radio BGI(k)	ELF(k)	ELP(k)
1	0.019	0.041	0.042	0.060	0.039	0.055	0.041	0.011	0.100	0.182
	(0.048)	(0.090)	(0.102)	(0.135)	(0.085)	(0.125)	(0.095)	(0.026)	(0.172)	(0.301)
2	0.037	0.073	0.078	0.110	0.070	0.089	0.080	0.021	0.180	0.318
	(0.070)	(0.119)	(0.135)	(0.179)	(0.121)	(0.161)	(0.137)	(0.037)	(0.206)	(0.342)
3	0.047	0.085	0.089	0.123	0.087	0.099	0.095	0.025	0.204	0.336
	(0.082)	(0.129)	(0.142)	(0.187)	(0.144)	(0.168)	(0.149)	(0.042)	(0.224)	(0.335)
4	0.058	0.098	0.102	0.148	0.104	0.119	0.117	0.029	0.246	0.379
	(0.088)	(0.134)	(0.147)	(0.196)	(0.155)	(0.175)	(0.163)	(0.044)	(0.242)	(0.330)
5	0.066	0.110	0.114	0.167	0.128	0.138	0.136	0.034	0.286	0.408
	(0.092)	(0.139)	(0.150)	(0.201)	(0.174)	(0.184)	(0.176)	(0.047)	(0.262)	(0.330)
6	0.076	0.128	0.129	0.190	0.148	0.159	0.157	0.040	0.319	0.415
	(0.101)	(0.152)	(0.159)	(0.213)	(0.191)	(0.199)	(0.194)	(0.052)	(0.284)	(0.323)
7	0.083	0.141	0.142	0.208	0.165	0.177	0.175	0.044	0.355	0.438
	(0.104)	(0.164)	(0.167)	(0.223)	(0.207)	(0.212)	(0.213)	(0.054)	(0.295)	(0.322)
8	0.084	0.147	0.147	0.216	0.172	0.185	0.185	0.046	0.369	0.447
	(0.104)	(0.166)	(0.170)	(0.225)	(0.211)	(0.215)	(0.216)	(0.054)	(0.296)	(0.315)
9	0.093	0.181	0.169	0.258	0.205	0.228	0.219	0.056	0.437	0.536
	(0.101)	(0.169)	(0.176)	(0.233)	(0.213)	(0.227)	(0.220)	(0.055)	(0.267)	(0.269)
10	0.101	0.202	0.184	0.285	0.231	0.254	0.243	0.064	0.493	0.561
	(0.099)	(0.176)	(0.180)	(0.234)	(0.219)	(0.231)	(0.224)	(0.060)	(0.264)	(0.248)
11	0.107	0.216	0.192	0.301	0.249	0.272	0.256	0.070	0.522	0.561
	(0.101)	(0.187)	(0.183)	(0.241)	(0.229)	(0.240)	(0.230)	(0.067)	(0.266)	(0.240)
12	0.107	0.216	0.193	0.301	0.250	0.272	0.257	0.071	0.523	0.560
	(0.101)	(0.188)	(0.183)	(0.242)	(0.229)	(0.240)	(0.231)	(0.067)	(0.267)	(0.240)
13	0.107	0.217	0.193	0.302	0.250	0.273	0.257	0.071	0.523	0.559
	(0.101)	(0.188)	(0.183)	(0.242)	(0.229)	(0.240)	(0.231)	(0.067)	(0.267)	(0.240)
Gini	0.490	0.678	0.679	0.912	0.734	0.800	0.805	0.438		
	(0.239)	(0.293)	(0.283)	(0.169)	(0.288)	(0.275)	(0.313)	(0.171)		
N	391	391	391	391	391	391	391	391	391	391
Primary	378	391	372	361	378	391	382	372	391	391

Notes. Mean values are reported for diversity indices, with respective standard deviations in parentheses. Overall Gini indices are provided for indicators in the corresponding column titles. The last row reports the number of observations covered by primary data sources.

Table D.3: Pairwise correlation coefficients for various BGI indices

		Education BGI				Electricity BGI				Television BGI				Car BGI			
		(1)	(5)	(9)	(13)	(1)	(5)	(9)	(13)	(1)	(5)	(9)	(13)	(1)	(5)	(9)	(13)
Education BGI	(1)	1															
	(5)	0.67	1														
	(9)	0.56	0.92	1													
	(13)	0.52	0.87	0.97	1												
Electricity BGI	(1)	0.53	0.31	0.24	0.22	1											
	(5)	0.42	0.62	0.58	0.53	0.69	1										
	(9)	0.27	0.40	0.50	0.51	0.59	0.79	1									
	(13)	0.20	0.27	0.39	0.46	0.51	0.63	0.91	1								
Television BGI	(1)	0.44	0.23	0.18	0.18	0.76	0.52	0.46	0.45	1							
	(5)	0.37	0.61	0.57	0.53	0.55	0.82	0.64	0.54	0.70	1						
	(9)	0.26	0.47	0.53	0.52	0.49	0.70	0.72	0.65	0.62	0.86	1					
	(13)	0.21	0.39	0.46	0.49	0.43	0.60	0.67	0.68	0.57	0.76	0.94	1				
Car BGI	(1)	0.30	0.13	0.09	0.10	0.61	0.40	0.36	0.35	0.71	0.49	0.43	0.39	1			
	(5)	0.29	0.53	0.51	0.48	0.41	0.64	0.50	0.40	0.47	0.76	0.65	0.56	0.63	1		
	(9)	0.16	0.37	0.45	0.48	0.33	0.50	0.64	0.61	0.34	0.56	0.61	0.56	0.51	0.80	1	
	(13)	0.12	0.27	0.37	0.45	0.27	0.38	0.59	0.64	0.29	0.45	0.53	0.57	0.44	0.67	0.93	1

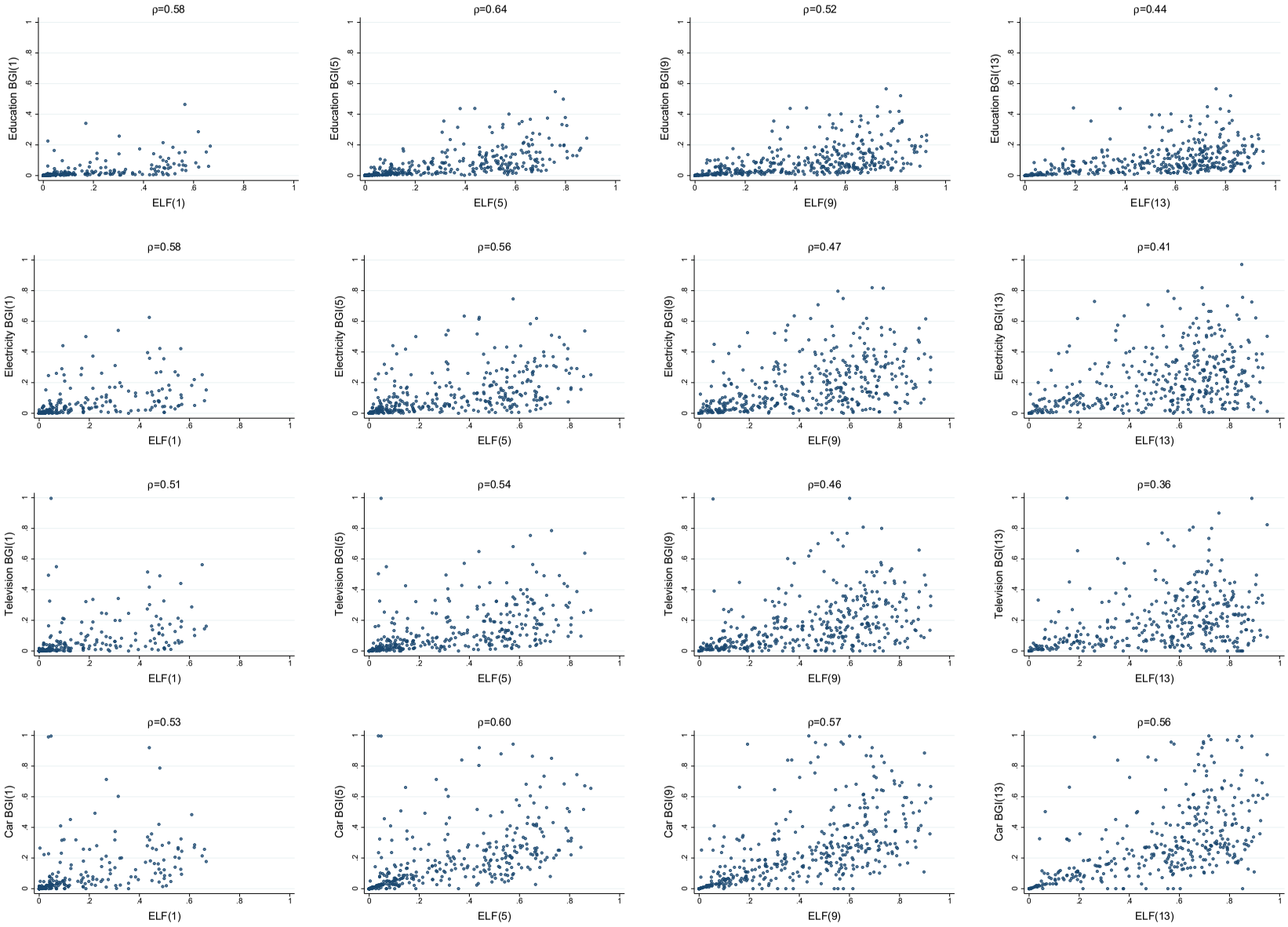
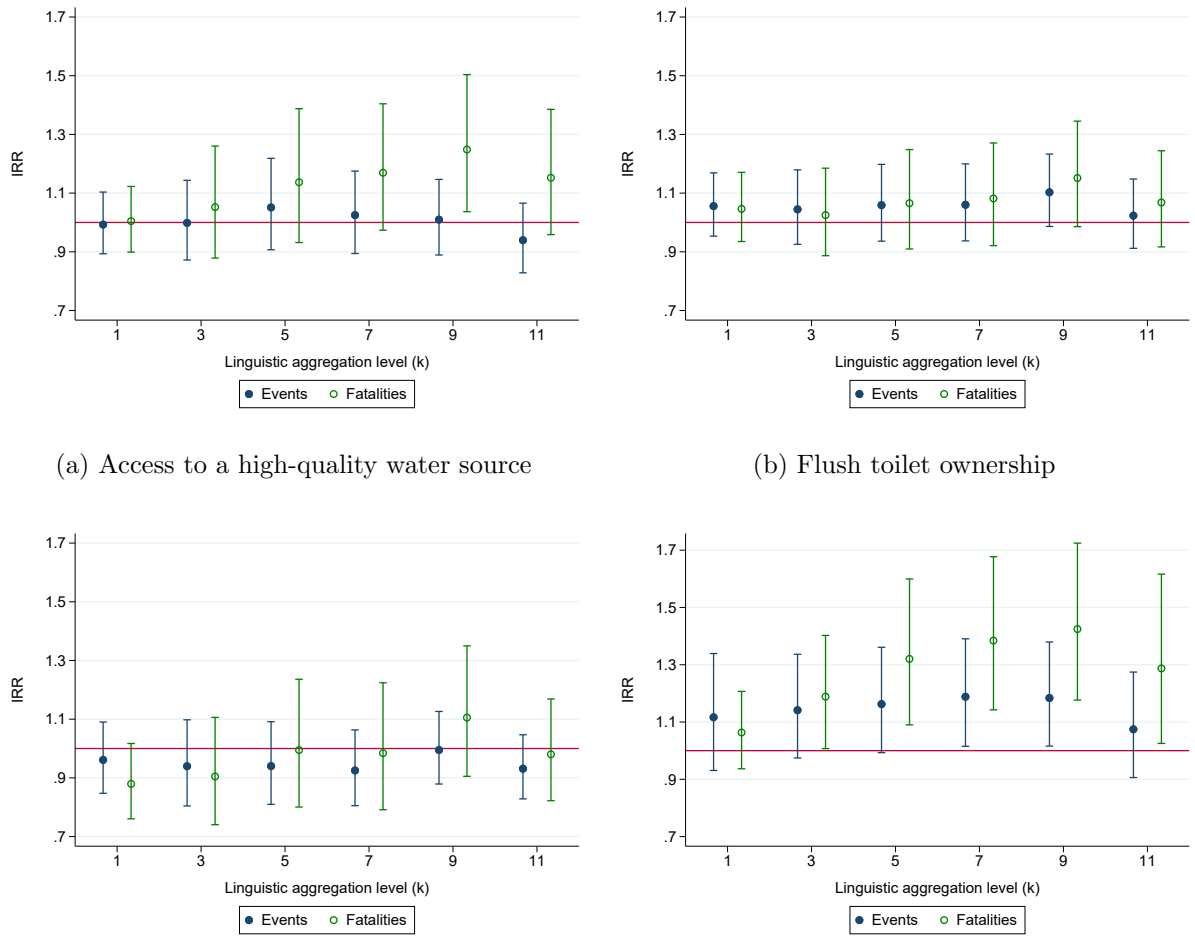


Figure D.1: Scatterplots and pairwise correlation coefficients (ρ) for selected ELF(k) and BGI(k) indices.

E Additional regression results

E.1 Further ethnic inequality indices and conflict



(a) Access to a high-quality water source

(b) Flush toilet ownership

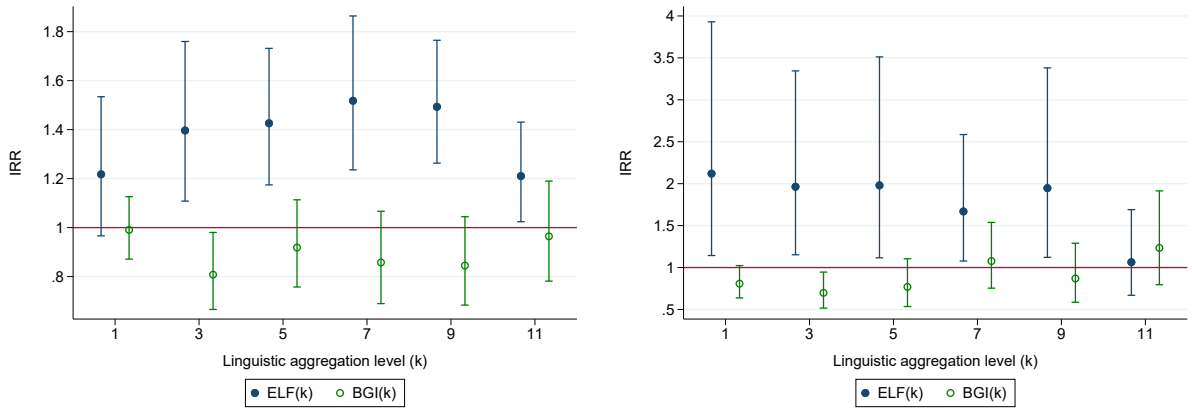
(c) Finished floor ownership

(d) Radio ownership

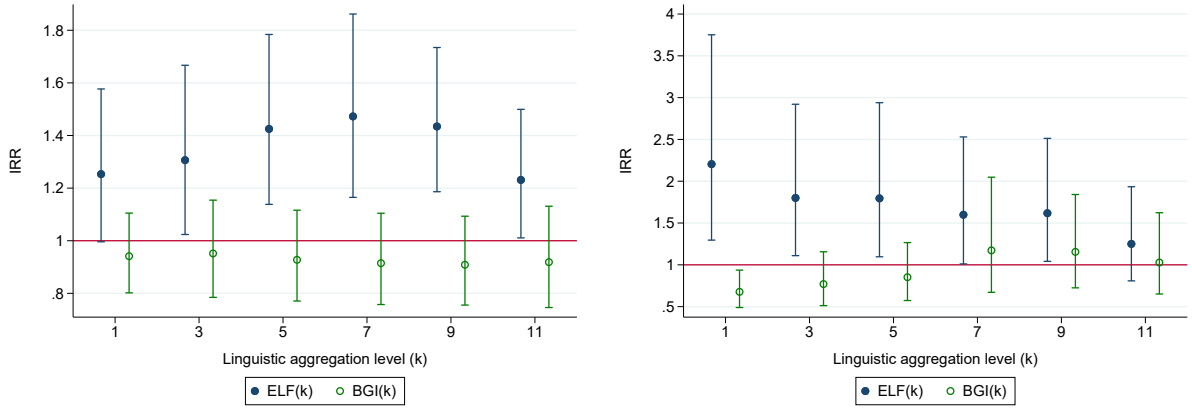
Figure E.1: Subnational ethnic inequality and conflict: additional BGI(k) measures.

Notes. Each panel presents incidence rate ratios, along with 95% confidence intervals, based on robust standard errors. For each of the six reported levels of linguistic aggregation, a negative binomial regression is estimated, where the outcome is the sum of either conflict events or fatalities, and the right-hand-side variable of interest is a BGI(k) index capturing ethnic inequality in the dimension indicated in the figure subtitle. All regressions include country fixed effects and baseline controls described in the main text. The number of observations is 391.

E.2 ELF versus BGI: Poisson regressions



(a) BGI in years of education and ELF as predictors of conflict events (left) and fatalities (right)

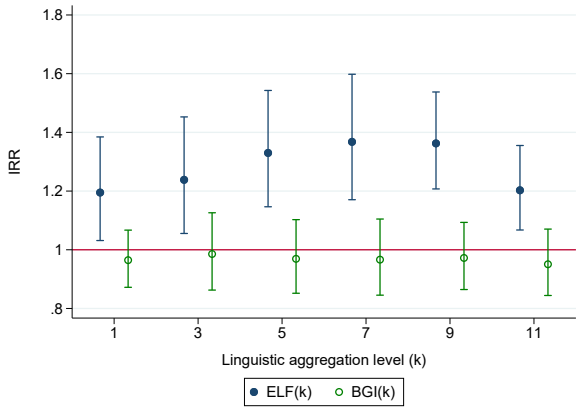


(b) BGI in electricity access and ELF as predictors of conflict events (left) and fatalities (right)

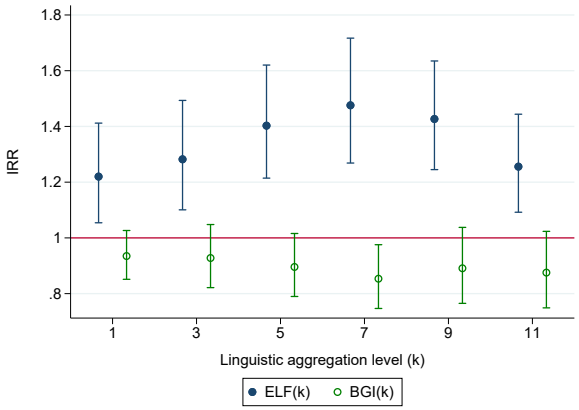
Figure E.2: Poisson "horse race" regressions for ELF and BGI.

Notes. Each panel presents incidence rate ratios, along with 95% confidence intervals, based on robust standard errors. For each of the six reported levels of linguistic aggregation, a Poisson regression is estimated, where the outcome is the sum of conflict events or fatalities. The right-hand side includes both ELF(k) and BGI(k) for years of education (panel a) electricity access (panel b), along with country fixed effects and baseline controls described in the main text. The number of observations is 391.

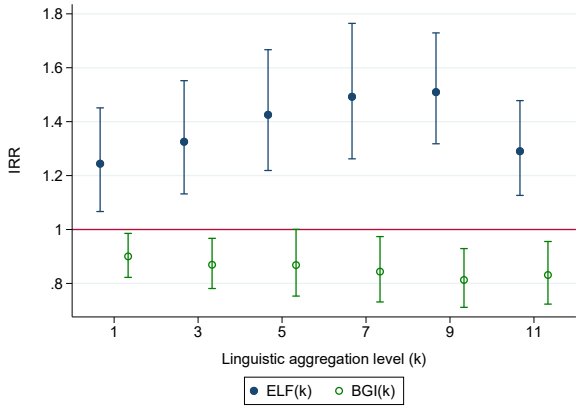
E.3 ELF versus BGI: Further "horse race" regressions



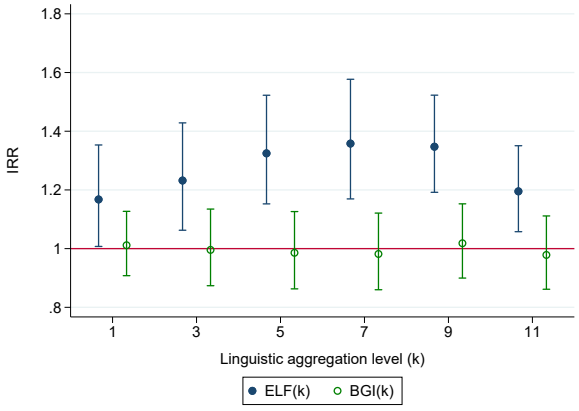
(a) Television ownership



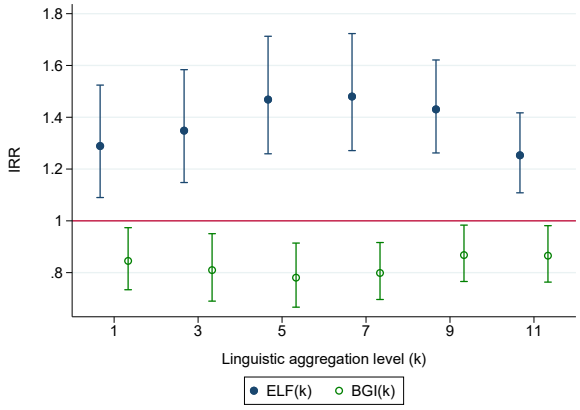
(b) Car ownership



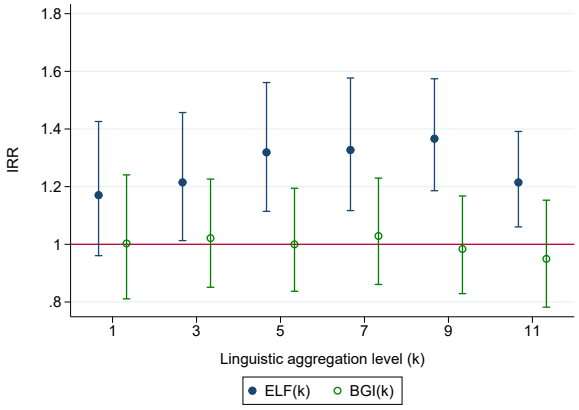
(c) Access to a high-quality water source



(d) Flush toilet ownership



(e) Finished floor ownership



(f) Radio ownership

Figure E.3: "Horse race" regressions for additional BGI(k) measures.

Notes. See figure 6 notes. The outcome variable is the count of conflict events.

References

Alcorta, Ludovico, Jeroen Smits, and Haley J. Swedlund, "Inequality and Ethnic Conflict in Sub-Saharan Africa," *Social Forces*, 2018, *97* (2), 769–792.

Alesina, Alberto and Eliana La Ferrara, "Ethnic Diversity and Economic Performance," *Journal of Economic Literature*, 2005, *43* (3), 762–800.

_ , Stelios Michalopoulos, and Elias Papaioannou, "Ethnic Inequality," *Journal of Political Economy*, 2016, *124* (2), 428–488.

Arbatlı, Cemal Eren, Quamrul H. Ashraf, Oded Galor, and Marc Klemp, "Diversity and Conflict," *Econometrica*, 2020, *88* (2), 727–797.

Baldwin, Kate and John D. Huber, "Economic Versus Cultural Differences: Forms of Ethnic Diversity and Public Goods Provision," *American Political Science Review*, 2010, pp. 644–662.

Berman, Nicolas, Mathieu Couttenier, Dominic Rohner, and Mathias Thoenig, "This Mine is Mine! How Minerals Fuel Conflicts in Africa," *American Economic Review*, 2017, *107* (6), 1564–1610.

Cederman, Lars-Erik, Nils B. Weidmann, and Kristian Skrede Gleditsch, "Horizontal Inequalities and Ethnonationalist Civil War: A Global Comparison," *American Political Science Review*, 2011, pp. 478–495.

Cervellati, Matteo, Giorgio Chiovelli, and Elena Esposito, "Bite and Divide: Malaria and Ethnolinguistic Diversity," 2019. CEPR Discussion Paper No. DP13437.

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

Davies, Shawn, Garoun Engström, Therése Pettersson, and Magnus Öberg, "Organized Violence 1989–2023, and the Prevalence of Organized Crime Groups," *Journal of Peace Research*, 2024, *61* (4), 673–693.

de la Cuesta, Brandon and Leonard Wantchekon, "Is Language Destiny? The Origins and Consequences of Ethnolinguistic Diversity of Sub-Saharan Africa," in Victor Ginsburgh and Shlomo Weber, eds., *The Palgrave Handbook of Economics and Language*, Palgrave Macmillan UK, 2016, chapter 18, pp. 513–537.

Desmet, Klaus, Ignacio Ortuño-Ortín, and Ömer Özak, "Is Secessionism Mostly About Income or Identity? A Global Analysis of 3,153 Subnational Regions," *Economic Journal*, 2025, *forthcoming.*

_ , _ , and Romain Wacziarg, "The Political Economy of Linguistic Cleavages," *Journal of Development Economics*, 2012, *97* (2), 322–338.

Esteban, Joan and Debraj Ray, "On the Salience of Ethnic Conflict," *American Economic Review*, December 2008, *98* (5), 2185–2202.

_ , Laura Mayoral, and Debraj Ray, "Ethnicity and Conflict: An Empirical Study," *American Economic Review*, June 2012, *102* (4), 1310–42.

Fearon, James D. and David D. Laitin, "Ethnicity, Insurgency, and Civil War," *American Political Science Review*, 2003, pp. 75–90.

Filmer, Deon and Kinnon Scott, "Assessing Asset Indices," *Demography*, 2012, *49* (1), 359–392.

Fjelde, Hanne and Gudrun Østby, "Socioeconomic Inequality and Communal Conflict: A Disaggregated Analysis of Sub-Saharan Africa, 1990–2008," *International Interactions*, 2014, *40* (5), 737–762.

Galor, Oded and Ömer Özak, "The Agricultural Origins of Time Preference," *American Economic Review*, 2016, *106* (10), 3064–3103.

Gershman, Boris and Diego Rivera, "Subnational Diversity in Sub-Saharan Africa: Insights From a New Dataset," *Journal of Development Economics*, 2018, *133*, 231–263.

_ and _ , "Measuring Regional Ethnolinguistic Diversity in Sub-Saharan Africa: Surveys vs. GIS," *The World Bank Economic Review*, 2020, *34* (S1), S40–S45.

Gibson, John, Susan Olivia, and Geua Boe-Gibson, "Night Lights In Economics: Sources and Uses," *Journal of Economic Surveys*, 2020, *34* (5), 955–980.

Ginsburgh, Victor and Shlomo Weber, "The Economics of Language," *Journal of Economic Literature*, June 2020, *58* (2), 348–404.

Hillesund, Solveig, Karim Bahgat, Gray Barrett, Kendra Dupuy, Scott Gates, Håvard Mokleiv Nygård, Siri Aas Rustad, Håvard Strand, Henrik Urdal, and Gudrun Østby, "Horizontal Inequality and Armed Conflict: A Comprehensive Literature Review," *Canadian Journal of Development Studies*, 2018, *39* (4), 463–480.

Hodler, Roland, Sorawoot Srisuma, Alberto Vesperoni, and Noémie Zurlinden, "Measuring Ethnic Stratification and Its Effect on Trust in Africa," *Journal of Development Economics*, 2020, *146*, 102475.

Houle, Christian and Cristina Bodea, "Ethnic Inequality and Coups in Sub-Saharan Africa," *Journal of Peace Research*, 2017, *54* (3), 382–396.

Howe, Laura D., James R. Hargreaves, and Sharon R.A. Huttly, "Issues in the Construction of Wealth Indices for the Measurement of Socio-Economic Position in Low-Income Countries," *Emerging Themes in Epidemiology*, 2008, *5* (3), 1–14.

Huber, John D. and Laura Mayoral, "Group Inequality and the Severity of Civil Conflict," *Journal of Economic Growth*, 2019, *24* (1), 1–41.

Kiszewski, Anthony, Andrew Mellinger, Andrew Spielman, Pia Malaney, Sonia Ehrlich Sachs, and Jeffrey Sachs, "A Global Index Representing the Stability of Malaria Transmission," *American Journal of Tropical Medicine and Hygiene*, 2004, *70* (5), 486–498.

Kuhn, Patrick M. and Nils B. Weidmann, "Unequal We Fight: Between- and Within-Group Inequality and Ethnic Civil War," *Political Science Research and Methods*, 2015, *3* (3), 543–568.

Leipziger, Lasse Egendal, "Measuring Ethnic Inequality: An Assessment of Extant Cross-National Indices," *British Journal of Political Science*, 2023, *53* (2), 652–673.

Lewis, M. Paul, ed., *Ethnologue: Languages of the World*, 16 ed., Dallas: SIL International, 2009.

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

Michalopoulos, Stelios, "The Origins of Ethnolinguistic Diversity," *American Economic Review*, 2012, *102* (4), 1508–39.

_ **and Elias Papaioannou**, "The Long-Run Effects of the Scramble for Africa," *American Economic Review*, 2016, *106* (7), 1802–48.

Mitton, Todd, "The Wealth of Subnations: Geography, Institutions, and Within-Country Development," *Journal of Development Economics*, 2016, *118*, 88–111.

Montalvo, José G. and Marta Reynal-Querol, "Ethnic Polarization, Potential Conflict, and Civil Wars," *American Economic Review*, 2005, *95* (3), 796–816.

_ **and** _ , "Cultural Diversity, Conflict, and Economic Development," in Victor A. Ginsburgh and David Throsby, eds., *Handbook of the Economics of Art and Culture*, Vol. 2, Elsevier, 2014, chapter 18, pp. 485–506.

Nunn, Nathan and Diego Puga, "Ruggedness: The Blessing of Bad Geography in Africa," *Review of Economics and Statistics*, February 2012, *94* (1), 20–36.

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

Posner, Daniel N., "Measuring Ethnic Fractionalization in Africa," *American Journal of Political Science*, 2004, *48* (4), 849–863.

Raleigh, Clionadh, Andrew Linke, Håvard Hegre, and Joakim Karlsen, "Introducing ACLED: An Armed Conflict Location and Event Dataset," *Journal of Peace Research*, 2010, *47* (5), 651–660.

Ray, Debraj and Joan Esteban, "Conflict and Development," *Annual Review of Economics*, 2017, *9*, 263–293.

Salehyan, Idean, Cullen S. Hendrix, Jesse Hamner, Christina Case, Christopher Linebarger, Emily Stull, and Jennifer Williams, "Social Conflict in Africa: A New Database," *International Interactions*, 2012, *38* (4), 503–511.

Smits, Jeroen and Roel Steendijk, "The International Wealth Index (IWI)," *Social Indicators Research*, 2015, *122* (1), 65–85.

Stewart, Frances, ed., *Horizontal Inequalities and Conflict: Understanding Group Violence in Multiethnic Societies*, Houndmills, UK: Palgrave Macmillan, 2008.

Sundberg, Ralph and Erik Melander, "Introducing the UCDP Georeferenced Event Dataset," *Journal of Peace Research*, 2013, *50* (4), 523–532.

Vogt, Manuel, Nils-Christian Bormann, Seraina Rüegger, Lars-Erik Cederman, Philipp Hunziker, and Luc Girardin, "Integrating Data on Ethnicity, Geography, and Conflict: The Ethnic Power Relations Data Set Family," *Journal of Conflict Resolution*, 2015, *59* (7), 1327–1342.