Intersectional Inequality in Education in Africa, Asia, and the Americas

December 23, 2023

Abstract

Intersectional inequality — the notion that disparities run along combinations of social groups such as gender or ethnicity — has become an increasingly prominent concept in the social sciences. However, there is little empirical research applying an intersectional framework to measure inequality. We propose two novel metrics of intersectional inequality based on the concept of horizontal inequality. Based on these measures, we analyze educational intersectionality in gender and ethnicity using data from 40 countries. We show that the intersectional perspective unveils substantial inequality that remains masked if gender and ethnicity are analyzed in isolation and which is more than the sum of gender and ethnic inequality. Reducing inequalities based on gender and ethnicity separately might not be enough to "leave no one behind."

JEL codes: D63, I24, J16

Keywords: Inequality, Intersectionality, Measurement, Education

1 Introduction

Leave no one behind, the central principle of the 2030 Agenda for Sustainable Development, highlights that "barriers people face in accessing services, resources and equal opportunities are not simply accidents of fate or a lack of availability of resources, but rather the result of discriminatory laws, policies and social practices that leave particular groups of people further and further behind" (UNSDG, 2022). In other words, disadvantages occur not only for individuals but also for whole social groups, for instance, women or members of marginalized ethnic groups. Since individuals have little (or no) influence over group membership, like gender or ethnicity, these systematic disparities are not only problematic from the perspective of equality of opportunity, but they are detrimental to economic development on a broader scale (Ferreira et al., 2018; Marrero and Rodríguez, 2013). In response, economists increasingly apply the concept of horizontal inequality when measuring inequalities between social groups, such as gender or ethnicity (Mancini et al., 2008).

At the same time, "intersectionality" has been intensively discussed in the social sciences. For example, the UN Sustainable Development Group states that identifying inequalities requires disaggregation beyond gender, geography, and age and should occur in multiple and intersecting ways (UNSDG, 2022). The term *intersectionality* was coined by Crenshaw (1989, 1991) as a critical theoretical framework to describe the distinct discrimination faced by members at the "intersection" of social groups. For example, Black women might face specific disadvantages that neither Black men nor white women experience. Similarly, Kabeer (2016) uses the term "intersecting inequalities" to highlight individuals' overlapping disadvantages, reinforcing their exclusion. The particular overlaps that characterize marginalization vary by context, but Kabeer (2016) points out that the most enduring forms of group-based disadvantages are strongly associated with identities (arguably) ascribed at birth, such as caste, gender, and ethnicity. Therefore, this framework could be viewed as an extension of the horizontal inequalities framework. However, intersectionality remains mostly exclusive to theory in the humanities and is only starting to gain traction in the quantitative social sciences. Moreover, there are very few examples directly linking the concept of intersectionality to group-based inequalities in low- and middle-income countries (see e.g., Kabeer and Santos, 2017; Lenhardt and Samman, 2015).

To fill this gap, our work seeks to reconcile the intersectionality framework with the measurement of horizontal inequalities. In particular, we propose two new measurements and analyze intersectional inequalities in educational attainment in 40 countries by combining gender and ethnicity to form intersecting groups. As long as educational outcomes differ systematically and substantially between social groups, it is unlikely

that the world will succeed in "leaving no one behind". We also focus on education for analytical reasons, since many other well-being indicators, such as income or wealth, are hardly separable among members of the same household. Thus, measuring gender inequalities in these outcomes is not feasible. Meanwhile, education is an outcome that accrues entirely to the individual, allowing us to analyze gender differences in schooling.

We analyze how educational attainment varies across intersecting groups and time. To this end, we combine data from several rounds of the Demographic and Health Surveys (DHS) in 39 low- and middle-income countries between 1992 and 2019, as well as data from the US Current Population Survey (CPS) 2019, resulting in 2,689,289 individual observations. Little is known about intersecting inequalities in the global context and the DHS data pose a unique opportunity to analyze this topic. We include US data since the intersectionality literature has its origins in the US, making it relevant to put the magnitude of the results for low- and middle-income countries into perspective.

Our first measure of intersectional inequality is the schooling ratio between the group with the lowest (most disadvantaged) and the group with the highest (most advantaged) average education. We do this across gender, ethnicity, and the combination thereof. Compared to other inequality measures, this approach emphasizes the extremes of the distribution, in line with the principle to "leave no one behind." We find that intersectional inequality between ethnicity and gender differs significantly across countries and is larger than horizontal inequality by ethnicity and gender separately. Intersectional inequality is mainly driven by ethnic inequality and less by gender inequality since ethnic inequalities still tend to be more pronounced in many countries. Interestingly, although the concept of intersectional inequality was coined in the US, we find some of the lowest intersectional inequalities in education in the US compared to other countries.

The second measure of intersectionality we propose — which we refer to as differential intersectionality — aims to quantify the intersectionality that is "more than the sum of its parts" of gender and ethnic inequality. To this end, we first estimate how much intersectional inequality would arise if gender inequality were constant across all ethnic groups as a synthetic counterfactual. We compare this measure with the total, first-order intersectional inequality and calculate the difference, yielding the differential intersectionality measure. In 27 countries, the intersecting inequality is positive, implying that it is greater than what would be expected if all ethnic groups experienced the same gender inequality, and thus "more than the sum of its parts". For 13 countries in the sample, differential intersectionality is negative, indicating less

gender inequality in the most advantaged/disadvantaged ethnic groups, or that the most disadvantaged group in a country are men of a particular ethnicity, despite in general women having lower levels of education than men.

We further use regression analysis to identify the correlation between the level of education and intersectionality. The analysis shows that while total intersectional inequality in education is associated with the general level of education, differential intersectionality is not. The number of ethnic groups in a country is also associated with higher intersectional inequality but not with differential inequality. Differences in sample size have a marginal effect on the measured level of intersectional inequality.

The remainder of this paper proceeds as follows. Section 2 discusses the related literature. Section 3 introduces the concept and two measures of intersectional inequalities and describes the empirical strategy to estimate intersectional inequalities and the subsequent analysis. Section 4 presents more information on the data. Section 5 presents the results of the analysis. Section 6 concludes.

2 Related Literature

This paper contributes to five different strands of literature touching on intersectional inequality.

First, on a broader level, this paper speaks to established theoretical literature in sociology, social psychology, and gender studies that conceptualizes intersectionality theoretically. Kimberlé Crenshaw (1989) coined the term, and what has followed is an ample discussion about the consequences of adopting an intersectional perspective, not only for social sciences but also for public policy (to name just a few examples: Alexander-Floyd, 2012; Berger and Guidroz, 2010; Bowleg, 2008; Cho et al., 2013; Choo and Ferree, 2010; Few-Demo, 2014; Hancock, 2007; Shields, 2008; Strid et al., 2013; Walby et al., 2012). However, the cited works make little to no prescriptions of *how* intersectionality could be operationalized quantitatively. The paper at hand contributes to this literature by proposing a framework for how researchers could include intersectionality in the quantitative measurement of inequality.

Second, the research that empirically applies an intersectional perspective primarily addresses education inequality at the intersection of race and gender in the US context. A number of studies examine the Black gender gap in college success (Keels, 2013; McDaniel et al., 2011; Mittleman, 2021), labor market returns to math performance (Riegle-Crumb, 2006), or success expectations in STEM-related subjects (Parker et al., 2020, review article). Together, the studies paint a clear picture of intersectional disparities in education. Black women do not fully profit from the generally closing

gender gap in education, yet Black men are typically even worse off than Black women. These effects are partly offset by socioeconomic status, especially for black women, implying that there is no gender gap for Black women with high socioeconomic status (Keels, 2013). These findings highlight the importance of intersectionality as an analytical framework because they provide essential insights usually lost when considering social identities like gender in isolation. However, the methodological frameworks applied in the cited studies are not necessarily applicable to other countries where the concepts of race and ethnic groups differ substantially from the US context.

For example, in the non-US context, Sen et al. (2009) analyze inequality in access to health services at the intersection of gender and social class in India. They find that the probability of non-treatment is only lower for women from poor households while being poor (or "lower class") is not relevant for men. The authors model intersectionality as interaction terms in their regressions and report the probability (odds ratio) of having a particular health-related outcome for members of a social group compared to a reference group. Our study departs from this literature in several aspects. On the one hand, we explicitly relate our empirical analysis to the theoretical literature on intersectionality. As such, intersectionality is not merely meant as an afterthought, but as an analytical lens through which to study inequalities. On the other hand, we emphasize the measurement of inequality as an outcome. Using inequality ratios, we express inequality in one measure instead of inferring differences between the groups from regression analyses. This method allows us to define intersectional inequality as a universal measure, irrespective of the specific context. Furthermore, we look beyond the US and focus on 39 low- and middle-income countries, which allows us to assess the relevance of intersectionality for a large part of the world's population and comparison across countries.

Third, our study integrates the concept of intersectionality into the growing literature on the measurement of horizontal inequality.[1] Shorrocks (1984) was among the first to propose the decomposition of inequality measures into population subgroups. This idea is mirrored in a growing body of research studying the concept of horizontal or between-group inequality (See e.g., Langer (2005), Langer et al. (2007), Mancini et al. (2008), Mancini (2008), Stewart (2009), Elbers et al. (2008), Cederman et al. (2011), Cederman et al. (2015), Canelas and Gisselquist (2018), Leivas and Dos Santos (2018), McDoom et al. (2019), Tetteh-Baah (2019)). To the best of our knowledge, only one study explicitly measures intersecting inequalities based on horizontal inequality across countries. Specifically, Lenhardt and Samman (2015) analyze intersecting inequalities

[1]For literature on inequality in its "vertical" sense, see e.g., Piketty and Saez (2014) for the US and Europe and Ravallion (2014) for developing countries.

in women's education (at the intersection of ethnicityand place of residence) in 16 low- and middle-income countries using DHS data, but do not include education data on men. Hence, this analysis leaves out differences in groups in the category of gender (see Figure 1).

Last, our research ties to a broader literature on gender inequality and ethnic and religious inequalities in education in the Global South. For gender inequality in education, see e.g., King and Hill (1995), Lopus and Frye (2018), Klasen (2002), Klasen and Lamanna (2009). For ethnic and religious inequalities, see e.g., Easterly and Levine (1997), Montalvo and Reynal-Querol (2003), Montalvo and Reynal-Querol (2005), Alesina et al. (2016), Houle and Bodea (2017), Muller (2017), Alcorta et al. (2018), Cooray and Potrafke (2011), Hajj and Panizza (2009). All of this literature documents large and persistent gaps in ethnic inequality and declining gender inequality over time.

3 Measures

3.1 Conceptual framework

The most common concept of inequality measurement is "vertical inequality" (Bourguignon, 1979; Cowell, 1988; Lambert and Aronson, 1993). Vertical inequality typically measures inequality between individuals within or across geographic or economic entities. Some measures, such as the Gini coefficient or the Theil index, take the whole distribution of an outcome into account, while others, such as the Palma Index or the P90/P10 ratio, compare specific percentiles of the distribution. In contrast to vertical inequalities, horizontal inequalities occur between different social groups of categories, such as gender, ethnicity, religion, or location. They are, thus, often referred to as between-group inequalities (see Figure 1).

Place Figure 1 here

For our proposed measures of intersectional inequalities, we closely follow the concept of horizontal inequality, but we add an additional dimension. As Figure 1 shows, instead of analyzing inequalities between genders and ethnicities separately (i.e., staying within the same category), we use "intersecting" groups, that is, we compare women and men who belong to different ethnic groups; hence, comparing across groups and categories (see Figure Figure 1). This intersectional perspective allows us to uncover gender differences within and between ethnic groups.[2]

[2]We are aware that a binary gender classification system might oversimplify the diversity of gender

3.2 Inequality ratio

To measure horizontal and intersectional inequality between social groups, we mainly focus on the inequality ratio (IR). It is a simple (unweighted) ratio between the group with the highest average of an outcome variable — in our case, years of education — and the group with the lowest average. Formally, we can describe IR in the following way. Let the mean in years of education of one group be defined as follows:

$$s_j = \frac{\sum^{i \in j} educ_i}{n_j}, \tag{1}$$

where n represents the number of observations in group j. Then, the inequality ratio IR for a type $k \in \{g, e, g \times e\}$, where $g = male, female$, and e for ethnicity, is calculated as follows:

$$IR(k) = 1 - \frac{min\{s_j, ..., s_J\}}{max\{s_j, ..., s_J\}}, \forall k \in \{g, e, g \times e\} \tag{2}$$

Because the numerator is weakly smaller than the denominator, the inequality ratio is bounded between 0 and 1. A value closer to 0 implies complete equality, whereas a value closer to 1 implies more inequality between the groups. In the event that one group has an average outcome of 0, the inequality ratio would be 1. In this situation, the ratio would not be affected by the average outcome of the most advantaged group. However, this is not a realistic scenario, as it is highly improbable that all members of a group have zero years of education and never occurs in the context of the data at hand.

Compared to other inequality measures, such as the Gini or Theil index, the inequality ratio mainly conveys information about the tails of the distribution (Conceicao and Ferreira, 2000; Cobham and Sumner, 2013) and is very intuitively interpretable: by subtracting its value from one, the ratio tells us what fraction of education the group with the lowest average education has compared to the group with the highest average education. For example, assuming that men have a higher average education than women (as is the case in most countries in our sample), an inequality ratio of $IR(gender) = 0.75$ means that women have, on average, 25% of the years of education of men. Moreover, when there are only two groups — as is the case for gender — there are hardly any reasons to resort to a more complex measure of horizontal inequality, such as the Gini index. When there are more than two groups, the inequality ratio

identities, including non-binary individuals. This oversimplification could result in an inadequate assessment of intersectional inequalities among the most marginalized groups. Unfortunately, the DHS datasets do not provide information on additional gender identities. We recognize the need for future surveys to address this issue by collecting data on a wider range of gender identities. This would enable more nuanced categorizations and facilitate more comprehensive analyses.

has the property of omitting large parts of the information by only comparing the two most extreme groups. However, considering the principle of "leave no one behind," one could argue that any observed differences between any social groups in the middle of the welfare distribution are irrelevant and that we are particularly interested in the group "left behind." Moreover, the concept of intersectionality is originally concerned with the most disadvantaged and not about the overall inequality between groups. Proponents of intersectionality highlight that by focusing on group inequality without studying the intersection of social groups, the most disadvantaged group might be overlooked.

What distinguishes the approach described in this paper from the standard measurement of horizontal inequalities is the intersectional perspective. In the measurement of standard horizontal inequalities, the groups j either represent one of two genders or one of several ethnic groups. Here, we use intersecting groups, that is, every combination of gender and ethnicity. In this simple setup, the number of groups is doubled compared to the original number of ethnic groups. That is, for each ethnic group, there is now a separate group for women and men.

We deliberately do not weigh the group averages by the corresponding population size. The motivation for this approach is that we do not want to impose relative importance on any group's outcome. On the contrary, we are particularly interested in the outcomes of minority groups. Our measure only considers the two extreme groups (most advantaged/disadvantaged), which in many cases represent minorities. Using population weights would lead to strong distortions if one of the extreme groups is particularly small while the other is large.

While we aim to ensure that the specific identity of a group achieving a particular education outcome does not influence the measure, it is important to note that our approach does not satisfy the traditional anonymity axiom of inequality measures at the individual level, because we take into account which group an individual belongs to. Instead, our measure is designed to emphasize the outcomes of the most and least advantaged groups without being influenced by the specific identity of those groups. Hence, we adhere to the anonymity axiom of inequality measures, in the sense that which group has a given (education) outcome does not matter. Focusing on the tails of the distribution, the measure does not fulfill the Pigou-Dalton-Transfer principle across the entire distribution. Being a ratio measure, it fulfills the axiom of scale independence often applied to the construction of inequality measures.

Counterfactual and differential intersectionality. We first compare the estimates of "first-order" total intersectional inequality ($gender \times ethnicity$) with the horizontal

inequality estimates based on single types (*gender* or *ethnicity*). However, these estimates only give us a limited benchmark to evaluate the relative importance of intersectional inequality. The issue with directly comparing the inequality ratio $IR(gender \times ethnicity)$ to $IR(gender)$ and $IR(ethnicity)$ is that part of the gap between the intersectional and the horizontal inequality measures based on single groups arises by design, or "mechanically." In other words, if there is at least some gender inequality within the most and least disadvantaged ethnic groups, one will always obtain a greater inequality ratio for the intersectional types relative to gender and ethnicity in isolation.

To avoid this problem, we could think of a hypothetical situation where the gender gap in education was constant between ethnic groups to avoid this problem. In other words, one can calculate the counterfactual ("mechanical") component of the intersectional inequality ratio by applying the same relative difference between women and men as measured for the entire country for the lowest and the highest educated ethnic group. Going forward, we refer to this measure as "counterfactual intersectional inequality." We then calculate the difference between the counterfactual intersectional inequality and the total intersectional inequality. As a result, we obtain a measure of "differential intersectionality." The higher the value of this measure, the larger the *additional* component of intersectional inequality, that is, the more a particular intersecting group is disadvantaged or advantaged relative to gender and ethnic inequality. In other words, intersectional inequality is larger than the sum of its parts (i.e., horizontal inequality) when differential intersectionality is positive.

Formally, we define counterfactual intersectional inequality in the following way. Let $s^k_{min} = min\{s_j, ..., s_J\}$ and $s^k_{max} = max\{s_j, ..., s_J\}$ denote the mean of the respective group with the lowest and highest education within a type k. Then the counterfactual gender differences within the two extreme ethnic groups can be written as

$$\begin{aligned} s'_{min} &= s^{k=e}_{min} \frac{s^{k=g}_{min}}{S_J}, \\ s'_{max} &= s^{k=e}_{max} \frac{s^{k=g}_{max}}{S_J}, \end{aligned} \tag{3}$$

where S_J denotes the overall (gender-weighted) average in education. Then the inequality ratio of the counterfactual intersectionality is

$$IR(g \times e)' = 1 - \frac{s'_{min}}{s'_{max}}. \tag{4}$$

This measure reveals what the inequality ratio would be if all ethnic groups had the same

level of gender inequality within their own group.[3] Differential intersectionality is just the difference between the total intersectional inequality ratio and the counterfactual intersectional inequality ratio:

$$\Delta_{gender \times ethnicity} = IR_{total}(g \times e) - IR(g \times e)' \tag{5}$$

Therefore, if $\Delta > 0$, it implies that the intersectional inequality is greater than what would be expected if gender inequality was the same for all ethnic groups.

4 Data

To construct the horizontal and intersectional inequality measures, we use data on individuals' education in years of schooling and measure inequality grouped by gender and ethnicity. From the total of 337 surveys from 84 countries of the DHS, we use all DHS rounds where both women and men were interviewed, and the respondent's gender, ethnicity, and years of education were recorded. The resulting sample of 39 low- and middle-income countries contains data from 1992 until 2019 (see Table A1 for a detailed list of countries and sample sizes per country). The sample consists predominantly of African countries, but South Asian and Latin American countries are also adequately represented. Together, the countries included in the sample account for roughly 24% of the world population and 80% of the population of all African countries. The sample is not globally representative of low- and middle-income countries due to the absence of Middle Eastern countries. This gap does not arise because there are no DHS data for this region, but because, in many cases, ethnicity was not elicited, or no data on men was administered as part of the DHS survey. Furthermore, China is not part of the DHS program at all, and for India, there is only data on caste membership, which is too distant from any definition of ethnicity to be comparable. In addition to the DHS data, we include data from the 2019 Current Population Survey (CPS) for the United States (US) (Flood et al., 2021). The reason for including the US is that the concept of intersectionality originates from the US context. Thus,it is relevant to see how global intersectional inequality in education based on ethnicity/race and gender compares to the US. In contrast, most European statistical offices do not record ethnicity, and therefore, this type of analysis cannot be conducted in this context.

We divide the data into three birth cohorts: up to 1969, between 1970 and 1979,

[3]When there are three or more categories that influence intersectionality, the calculation becomes more difficult. For example, one would need to keep inequality constant within one category, such as gender, and then apply the overall gender difference to the intersecting groups between the other two categories. The complexity increases with the number of categories included.

and 1980 and later. This is done to ensure that the countries have a considerable amount of overlap over time. By breaking the data into cohorts, we can also analyze educational trends across time. In total, this approach yields 97 cells of cohorts and countries. The choice of the cohort is mainly driven by the data availability of DHS countries and survey years. Since we are looking at years of education completed, we focus on adults to ensure full exposure to the regular school system. We limit the sample to respondents 25 years and older to mitigate the problem that many younger respondents could still be in school.[4]

Education, the primary variable of interest, is measured by years of schooling. We prefer years of schooling over, for example, the highest completed level of education because it is the most widely elicited statistic of education. Moreover, we set an upper limit of 17 years of education (12 years of primary and secondary education and five years of tertiary education).[5]

To calculate between-group inequality, we use gender and ethnicity as categories. While it would be conceivable to analyze other categories, such as religion, region, or urban vs. rural residence, we decided to focus on the two categories most unequivocally ascribed at birth. However, in contrast to gender, ethnicity poses two challenges when used as a grouping variable. First, ethnic groups are not harmonized across the DHS survey rounds. In other words, the definition of ethnic groups is not consistent between survey rounds within a given country. Second, many ethnic groups have a small sample size. Particularly for younger and older cohorts, the number of observations for a given cell (a cohort-ethnicity-gender combination) is too small for a meaningful statistical analysis. To counter these challenges, we harmonize ethnic groups across survey rounds to ensure a minimum of 40 observations per cell and consistent grouping. As a first step, we harmonize ethnic groups between survey rounds within countries. For this purpose, we identify the larger groups to which smaller subgroups belong if they only appear in particular DHS rounds. To identify connections, we rely primarily on the online database Ethnologue. In a second step, we count the observations for each combination of ethnic groups, gender, and cohort bracket. We then merge each ethnic subgroup where one gender does not reach at least 40 observations for a cohort bracket into a larger ethnic group, again using the Ethnologue database. When merging is impossible, smaller ethnic groups are lumped into a separate group labeled "other." Thus, because,

[4]For example, from a survey from 2020, only persons born before 1996 would be taken into account. In doing so, we end up with three cohorts, each capturing a period of around 10 to 15 years with similar sample sizes. Countries have experienced different historical events, such as violent conflict or the timing of the end of colonialism, which we cannot precisely address with our cohort choice. However, the choice of using the same cohorts across countries allows us to describe levels and trends in intersectional inequality that allow for a comparison of results across countries.

[5]People who hold advanced tertiary degrees, such as a Ph.D., are relegated to 17 years of education.

by construction, groups that were merged will always have less extreme values than separate ones, the resulting intersectional inequality ratio should be thought of as a lower bound.

See Table A3 in the appendix for a detailed list of the surveys used and the ethnic groupings.

5 Results

Table 1 shows descriptive statistics for key variables of our analysis divided by birth cohort brackets. Only countries with observations in all three cohort brackets were included in this table in order to ensure comparability over time. It shows an increase in education of 1.4 years between the first (column 1) and the third (column 3) cohort brackets and a decrease of 24 percent in general inequality, 11 percent in ethnic inequality, and 12 percent in intersectional inequality. The average number of ethnic groups remains constant, which is expected given that we harmonized ethnic groups across the survey rounds.

Place Table 1 here

5.1 Total intersectional inequality

Figure 2 shows the estimates for the inequality ratios by gender, ethnicity, and intersecting groups (*gender* × *ethnicity*). An inequality ratio of 0 implies perfect equality between the groups with the lowest and highest average education. On the contrary, a ratio close to 1 indicates larger disparities between the two most extreme groups.

Place Figure 2 here

The ratios for gender are usually low, suggesting that there is relative equality in education among the countries studied. However, some countries, such as Afghanistan, Niger, and Chad, still have a large gap in education between men and women. In Afghanistan, the country with the highest gender inequality ratio of 0.77, women typically have only 23% of the education that men have. The ratios for ethnicity are usually higher than those for gender, indicating a greater inequality between the most and least disadvantaged groups. Only a few countries have lower inequality between ethnic groups than gender (Afghanistan, Guinea, Togo, Liberia, and Kazakhstan). African countries have particularly high ethnic disparities compared to Asia and the Americas, with inequality ratios as high as 0.83 in Chad and 0.87 in Burkina Faso.

The inequality ratios between intersecting groups (*gender* × *ethnicity*) are shown in the lightest shade. Intersectional inequality is higher than horizontal inequality based on gender and ethnic groups, which is to be expected (see Section 3). Figure 2 also reveals that intersectionalities are very high. In almost half of the countries in our sample, intersectional inequality ratios are above 0.7. This implies that in almost half of the countries, women of one ethnic group, on average, have less than a third of the years of education than men of another ethnicity. For instance, for the most recent cohort (1980 and younger) in Ethiopia, *Nuer* men have an average of 10.01 years of education, while *Affar* women have an average of 0.35 years, resulting in an inequality ratio of 0.97 (Figure B1 displays the group means by gender, ethnic, and intersecting groups). Nevertheless, in eleven countries, for at least one cohort group, the gender order is reversed, which means that either the group with the lowest education is male or the group with the highest education is female (see Table A2 in Appendix A). This is the case in Albania, Brazil, Burkina Faso, Guyana, Honduras, Kazakstan, Moldova, Mozambique, the Philippines, South Africa, and the US.

Furthermore, there are five countries where the most disadvantaged intersecting group has a different gender than the generally more disadvantaged gender in at least one cohort group. In twelve countries, the ethnicity of the most disadvantaged intersecting group does not match the overall most disadvantaged ethnic group. An intersectional perspective reveals inequalities that would be overlooked if gender and ethnic inequality were examined independently.

In the United States, the most significant source of intersectional inequality is racial disparities rather than gender disparities. This is more pronounced than the ethnic inequalities in countries with a similar high average education in our sample. However, in total intersectional inequality in the US seems to be rather low despite the concept having its scientific origins in the US.

Analyzing intersectionalities over time, we find that the more disadvantaged gender is constant across cohorts in all countries (see Table A2 in the Appendix). In only six countries, the most disadvantaged ethnic group changes over time, whereas the most advantaged ethnic group changes in 16 countries over time. This indicates that disadvantages based on ethnicity are more persistent than privileges.

The Spearman correlation between intersecting and ethnic inequality is 0.95, while the correlation between intersecting and gender inequality is 0.83. This suggests that intersecting inequality is more strongly influenced by ethnicity than gender, with ethnic inequalities being more pronounced and variable across countries than gender inequalities. Figure 2 also supports this conclusion.

5.2 Differential intersectionality

As stated in Section 3, the intersectional inequality ratio will always be greater (thereby demonstrating higher inequality) than the horizontal inequality ratios for gender or ethnicity. This was evident in Figure 2: as long as there is some gender inequality in the most disadvantaged ethnic group, the most disadvantaged intersecting group will always have a lower average than the ethnic group with the lowest education. This property is *mechanical*, meaning that it is not a consequence of intersectionality in its narrow definition, i.e., being more than the sum of its parts of gender and ethnic inequality, but that it is generated *by design*. Therefore, we calculate a measure for *counterfactual intersectional inequality* that reflects the intersectional inequality that would exist if all ethnic groups had the same gender inequality as the entire population.

Figure 3 illustrates the contrast between counterfactual intersectional inequality and total intersectional inequality, which is referred to as *differential intersectionality*. A zero value implies that the total and counterfactual intersectional inequality are equal, meaning that there is no differential intersectionality. A value above zero implies that the total intersectional inequality is greater than the counterfactual inequality ratio, indicating that the inequality between groups is greater than what would be expected from combining ethnic and gender inequalities. Conversely, a value below zero implies that the total intersectional inequality is lower than the counterfactual intersectional inequality.

Place Figure 3 here

Twenty-seven out of 40 countries show positive differential intersectionality. In these countries, total intersectional inequality is higher than counterfactual intersectional inequality. In other words, the gender gap is particularly wide for the most advantaged or disadvantaged ethnic groups. In some sense, one could say that in these countries, measuring inequality in an intersectional manner highlights that women (sometimes men) of certain ethnic groups are particularly disadvantaged or men (sometimes women) of particular ethnic groups are particularly advantaged. For instance, in Brazil, the country with the highest differential intersectionality, the total intersectional inequality ratio is 0.34 while the counterfactual ratio is 0.27, resulting in a differential intersectionality of 0.07 (or seven percentage points). This difference is substantial considering the already high total and counterfactual intersectional inequality values.

Meanwhile, 13 out of 40 countries exhibit values below zero. This result can only arise if the gender gap is narrower for the most advantaged or the most disadvantaged ethnic groups compared to the overall population of that country or if the gender inequality is reversed for the most disadvantaged ethnic group: for example, if for a

country in general women show lower education levels than men, but for the most disadvantaged ethnic group, it is men who have a lower education than women. For example, in Zimbabwe, the country with the lowest differential intersectionality, one would expect a counterfactual inequality ratio of 0.66 if the relative gender gap was the same across all ethnic groups. However, the total intersectional inequality ratio is 0.58, resulting in a negative differential intersectionality of -0.08 or 8 percentage points. This implies that the observed gender gaps in the most extreme groups are smaller than for the overall population or even reversed. This is also the case for the US.

5.3 Simulated sensitivity analysis

As we have seen in Figure 2, there is considerable variation in intersectional inequality across countries from 0.11 (South Africa) to 0.97 (Afghanistan), which raises questions regarding the role of sample size as a driver of this variation. In principle, by the law of large numbers, neither the number of groups nor the relative group size should impact the inequality ratio if the sample size approaches infinity. In other words, if the sample is large enough, the inequality ratio between arbitrarily drawn partitions of the sample (i.e., groups) converges to zero: perfect equality across groups, no matter how many groups. However, in reality, sample sizes are limited, and with an increasing number of groups, "extreme" mean values for one group become more likely by chance because the number of observations per group becomes small. Hence, part of the intersectionalities we observe might be due to small sample sizes in combination with many groups.

Place Figure 4 here

We conduct a sensitivity analysis to test whether the sample size and the number of groups play a significant role in determining the measured inequality ratios in our sample. We simulate samples by randomly drawing education and group membership from a given distribution, varying sample sizes and the number of groups. By doing so, we can estimate how much of the total intersectional inequality across countries could be caused by differences in sample sizes and the number of groups. Most country-cohorts in our data set have between 5,000 and 20,000 observations. But a couple also have only 2,000 observations (see Table A1 in the Appendix). The number of groups (gender times ethnicity) ranges mainly from 6-20 intersecting gender-ethnicity combinations (see Table A3), but two countries have 36 groups (Zambia and Ethiopia). However, their respective sample sizes are 10,000-20,000. Of course, ultimately, the determining factor will be the group sizes of the two extreme groups, but in order to give a reference point to the reader, we present the results for combinations of

total sample sizes and number of groups instead. Since groups are equally sized, the resulting group size can be inferred from the sample size divided by the number of groups in the corresponding simulation.

Figure 4 shows that with up to 16 groups, differences in sample size have a negligible effect on the inequality ratio. The simulated inequality ratio by chance is one with four groups (the minimum number of groups for an intersectional inequality ratio with gender and ethnicity), whether there are 30,000 or only 2,000 observations. However, a small sample size in combination with more than 16 groups might lead to high measured inequality by chance. For example, with 32 groups and 2,000 observations (62 observations per group), the simulated inequality ratio is close to 0.25. In Ethiopia and Zambia, the number of intersecting groups is 36 (18 ethnic groups split by gender). However, the sample size for these countries is between 10,000 and 20,000. Hence, we could observe an inequality ratio of approximately 0.10 by chance, which would be a 10% upward bias (if the sample size was 15,000 and there were 32 groups, as seen in Figure 4). This value is a much lower intersectional inequality than we observe for these countries (0.97 for Ethiopia and 0.5 for Zambia), but still, a fact that we need to take into account. Hence, we analyze the impact of sample size and the number of ethnic groups (given that the number of observed genders is the same across countries) as well as the cell size of the most disadvanted and most privilged group on measured inequality ratios in the next section.

5.4 Correlates of intersectional inequality

In Table 2, we explore correlates of intersectional inequality in education. While the results do not identify the causal drivers of intersectional inequality, this correlational analysis is still useful to analyze time trends as well as geographical patterns in inequalities and whether general improvements in education are likely to benefit all groups equally, i.e., reduce intersectional inequalities. In Panel A we focus on levels in education as the main variable of interest. In Panel B we focus on the number of ethnic groups and the total sample size, and in Panel C on the group size of the most disadvantaged and most privileged group. We always control for time trends.

Place Table 2 here

Columns A and B in Panel A show that the average level of education in a given country is associated with less intersectional inequality in education. The same result is found when looking at the counterfactual intersectional inequality (Columns C and D). When looking at the differential intersectionality (difference between counterfactual and total intersectional inequality), we find no significant association (Columns E and

F). In addition, we also do not find any general global time trend on intersectional inequalities when controlling for changes in mean education, as can be seen from the cohort fixed effects.

Panel B shows a significant association between the number of ethnic groups and intersectional inequality. A higher number of ethnic groups is positively associated with both intersectional inequality and counterfactual intersectional inequality. Similar to the results of education, we found no significant association with differential inequality. As expected, a significant but very small negative association is found between the sample size and all forms of intersectional inequality (Columns A to F), indicating that they have a negligible effect on observed differences in intersectional inequalities across countries.

Panel C shows the results for the group size of the most disadvantaged and most priviledged groups. Adding the two variables of the group size of the lowest and the highest group only marginally changes the results obtained in Panel B. As expected, no significant association is found between group sizes and intersectional inequality and counterfactual intersectional inequality. A small but significant association is found between the group size of the lowest group and differential intersectional inequality, but this effect is very small, indicating that it also has a negligible effect on observed differences in intersectional inequalities across countries.

Interestingly, in Panel C, we find a significant Africa dummy, and in Panel B a significant time trend for intersectional inequalities. It appears that the African geographical context and time trends have a diminished impact on intersectional inequality when taking into account the average education in a country. This implies that the general low education level in African countries is the main factor contributing to higher intersectional inequality. However, this significance disappears when controlling for overall average education, suggesting that education levels mediate the relationship between the African geographical context and intersectional inequality.

We perform a robustness check for Column A by using a group Theil index as a dependent variable in Table C1. The results are largely similar.

6 Conclusion

In this study, we propose a new approach to incorporate the intersectionality framework into the measurement of horizontal inequalities based on extensive household survey data from 40 countries. Hence, we bridge the gap between more qualitatively oriented social sciences and their conceptualization of inequalities and quantitative fields of social sciences, which have so far focused on measuring vertical and, to some extent,

horizontal inequalities. We show that framing horizontal inequality from an intersectional perspective reveals large disparities between women and men of different ethnic groups. We show that in many cases, specific combinations of gender and ethnicity uncover disadvantaged groups that are not apparent when gender and ethnicity are analyzed separately. Moreover, we show that in about two-thirds of the sample, there is differential intersectionality, indicating more inequality than "the sum of its parts" (gender and ethnic inequality).

Furthermore, we find that the total intersectional inequality is correlated with the education level in a country and the number of ethnic groups. The situation is different for differential intersectionality, where general improvements in education or a decrease in the number of ethnic groups do not decrease the intersectional inequality that is more than the sum of gender and ethnic inequality. Sample size plays a negligible role in the measured differences in intersectional inequality across countries.

Given these results, policymakers should take intersectionalities into account when designing and implementing programs to increase access to education. In countries with high differential intersectionality, special attention should be paid to targeting education policies, as general increases in education levels and decreases in gender and ethnic inequalities might still leave the most vulnerable groups behind.

Yet, our analysis also faces some limitations that should be taken into account when interpreting our findings. First, the definition of the social identity that we summarize under the umbrella term "ethnicity" actually differs substantially across countries. In large parts of Africa, ethnic groups run along the lines of languages, dialects, or tribal kinship. In other contexts, such as the US or Brazil, the more salient social identity is defined by a mixture of skin color, national origin, and indigenous status. Thus, cross-country comparisons of ethnic group inequality should be made with caution. Nevertheless, it is worth reemphasizing that any large disparities in education between "arbitrarily" generated groups are likely to be an indicator of discrimination.

Second, in contrast to the Gini or Theil index, the inequality ratio only considers the two most extreme groups. Therefore, we cannot draw any conclusions about the middle of the education distribution. However, in our view, against the backdrop of the original concept of intersectionality and the international agenda of *leaving no one behind*, it is warranted to consider only the most disadvantaged groups.

Third, our inequality measures do not directly give us any information about the gender of the two most extreme intersecting groups. For instance, it could be that, in some cases, the most advantaged and disadvantaged intersecting groups are both of the same gender. For the sake of simplicity, we leave the resolution of this issue to future research.

Lastly, the measure of differential intersectionality does not distinguish whether intersectionality inequality occurs because one group is particularly advantaged or disadvantaged. Yet, the two cases do not have the same implications for policy making. While the former would imply that education policy should be spread as widely as possible, the latter calls for specifically targeted interventions, for example, for women of most disadvantaged ethnic groups. To avoid this problem, one could, for instance, contrast the relative disadvantage of the women in the lowest educated ethnic group with the country's average rather than compare it against the most privileged group. However, the drawback of this approach is that it would not reflect inequality arising due to elite capturing. Consequently, we leave this line of questioning to future research.

Another open question for future research is the cause of intersectional inequality in different contexts. What is the social environment, and what are the policies and institutional factors that shape intersectional inequalities in education in specific contexts? Due to limitations in sample size, we did not analyze this question in our analysis and more rigorous research is needed to fully answer this question. Our approach could also be extended to other social identity groups (such as religion or place of residence) and other well-being indicators, such as health or wealth.

References

Alcorta, L., Smits, J., and Swedlund, H. J. (2018). Inequality and ethnic conflict in Sub-Saharan Africa. *Social Forces*, 97(2):769–792. doi: 10.1093/sf/soy049.

Alesina, A., Michalopoulos, S., and Papaioannou, E. (2016). Ethnic inequality. *Journal of Political Economy*, 124(2):428–488. doi: 10.1086/685300.

Alexander-Floyd, N. G. (2012). Disappearing Acts: Reclaiming Intersectionality in the Social Sciences in a Post—Black Feminist Era. *Feminist Formations*, 24(1):1–25. http://www.jstor.org/stable/23275087.

Berger, M. T. and Guidroz, K. (2010). *The Intersectional Approach: Transforming the Academy through Race, Class, and Gender*. University of North Carolina Press.

Bourguignon, F. (1979). Decomposable Income Inequality Measures. *Econometrica*, 47(4):901. doi: 10.2307/1914138.

Bowleg, L. (2008). When Black + Lesbian + Woman \neq Black Lesbian Woman: The Methodological Challenges of Qualitative and Quantitative Intersectionality Research. *Sex Roles*, 59(5):312–325. doi: 10.1007/s11199-008-9400-z.

Canelas, C. and Gisselquist, R. M. (2018). Horizontal inequality as an outcome. *Oxford Development Studies*, 46(3):305–324. doi: 10.1080/13600818.2018.1508565.

Cederman, L. E., Weidmann, N. B., and Bormann, N. C. (2015). Triangulating horizontal inequality: Toward improved conflict analysis. *Journal of Peace Research*, 52(6):806–821. doi: 10.1177/0022343315597969.

Cederman, L. E., Weidmann, N. B., and Gleditsch, K. S. (2011). Horizontal inequalities and ethnonationalist civil war: A global comparison. *American Political Science Review*, 105(3):478–495. doi: 10.1017/S0003055411000207.

Cho, S., Crenshaw, K. W., and McCall, L. (2013). Toward a Field of Intersectionality Studies: Theory, Applications, and Praxis. *Signs: Journal of Women in Culture and Society*, 38(4):785–810. doi: 10.1086/669608.

Choo, H. Y. and Ferree, M. M. (2010). Practicing Intersectionality in Sociological Research: A Critical Analysis of Inclusions, Interactions, and Institutions in the Study of Inequalities. *Sociological Theory*, 28(2):129–149. doi: 10.1111/j.1467-9558.2010.01370.x.

Cobham, A. and Sumner, A. (2013). Is It All About the Tails? The Palma Measure of Income Inequality. Center for Global Development Working Paper No. 343. doi: 10.2139/ssrn.2366974.

Conceicao, P. and Ferreira, P. (2000). The Young Person's Guide to the Theil Index: Suggesting Intuitive Interpretations and Exploring Analytical Applications. UTIP Working Paper No. 14. doi: 10.2139/ssrn.228703.

Cooray, A. and Potrafke, N. (2011). Gender inequality in education: Political institutions or culture and religion? *European Journal of Political Economy*, 27(2):268–280. doi: 10.1016/j.ejpoleco.2010.08.004.

Cowell, F. A. (1988). Inequality Decomposition: Three Bad Measures. *Bulletin of Economic Research*, 40(4):309–312. doi: 10.1111/j.1467-8586.1988.tb00274.x.

Crenshaw, K. (1989). Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics. *University of Chicago Legal Forum*, 1989(1):8.

Crenshaw, K. (1991). Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color. *Stanford Law Review*, 43(6):1241. doi: 10.2307/1229039.

Easterly, W. and Levine, R. (1997). Africa's growth tragedy: Policies and ethnic divisions. *Quarterly Journal of Economics*, 112(4):1203–1250. doi: 10.1162/003355300555466.

Elbers, C., Lanjouw, P., Mistiaen, J. A., and Özler, B. (2008). Reinterpreting between-group inequality. *Journal of Economic Inequality*, 6(3):231–245. doi: 10.1007/s10888-007-9064-x.

Ferreira, F. H. G., Lakner, C., Lugo, M. A., and Özler, B. (2018). Inequality of Opportunity and Economic Growth: How Much Can Cross-Country Regressions Really Tell Us? *Review of Income and Wealth*, 64(4):800–827. doi: 10.1111/roiw.12311.

Few-Demo, A. L. (2014). Intersectionality as the "New" Critical Approach in Feminist Family Studies: Evolving Racial/Ethnic Feminisms and Critical Race Theories. *Journal of Family Theory & Review*, 6(2):169–183. doi: 10.1111/jftr.12039.

Flood, S., King, M., Rodgers, S., Ruggles, S., Warren, and Westberry, W. (2021). Integrated Public Use Microdata Series, Current Population Survey: Version 9.0 [dataset]. Minneapolis, MN: IPUMS. doi: 10.18128/D030.V11.0.

Hajj, M. and Panizza, U. (2009). Religion and education gender gap: Are Muslims different? *Economics of Education Review*, 28(3):337–344. doi: 10.1016/j.econedurev.2008.01.007.

Hancock, A.-M. (2007). When Multiplication Doesn't Equal Quick Addition: Examining Intersectionality as a Research Paradigm. *Perspectives on Politics*, 5(1):63–79. doi: 10.1017/S1537592707070065.

Houle, C. and Bodea, C. (2017). Ethnic inequality and coups in sub-Saharan Africa. *Journal of Peace Research*, 54(3):382–396.

Kabeer, N. (2016). 'Leaving no one behind': the challenge of intersecting inequalitites. In *World Social Science Report 2016, Challenging Inequalities: Pathways to a Just World*. UNESCO Publishing, Paris. 55-8.

Kabeer, N. and Santos, R. (2017). Intersecting inequalities and the Sustainable Development Goals: Insights from Brazil. Technical Report 167, United Nations University World Institute for Development Economics Research.

Keels, M. (2013). Getting Them Enrolled Is Only Half the Battle: College Success as a Function of Race or Ethnicity, Gender, and Class. *American Journal of Orthopsychiatry*, 83(2,3):310–322. doi: 10.1111/ajop.12033.

King, E. and Hill, M. A. (1995). Women's education and economic well-being. *Feminist Economics*, 1(2):21–46. doi: 10.1080/714042230.

Klasen, S. (2002). Low schooling for girls, slower growth for all? cross-country evidence on the effect of gender inequality in education on economic development. *World Bank Economic Review*, 16(3):345–373. doi: 10.1093/wber/lhf004.

Klasen, S. and Lamanna, F. (2009). The impact of gender inequality in education and employment on economic growth: New evidence for a panel of countries. *Feminist Economics*, 15(3):91–132. doi: 10.1080/13545700902893106.

Lambert, P. J. and Aronson, J. R. (1993). Inequality Decomposition Analysis and the Gini Coefficient Revisited. *The Economic Journal*, 103(420):1221. doi: 10.2307/2234247.

Langer, A. (2005). Horizontal inequalities and violent group mobilization in Côte d'Ivoire. *Oxford Development Studies*, 33(1):25–45. doi: 10.1080/13600810500099634.

Langer, A., Mustapha, A. R., and Stewart, F. (2007). Horizontal Inequalities in Nigeria, Ghana and Côte d'Ivoire: Issues and Policies. CRISE Working Paper No. 45.

Leivas, P. H. S. and Dos Santos, A. M. A. (2018). Horizontal inequality and ethnic diversity in brazil: Patterns, trends, and their impacts on institutions. *Oxford Development Studies*, 46(3):348–362. doi: 10.1080/13600818.2017.1394450.

Lenhardt, A. and Samman, E. (2015). In quest of inclusive progress Exploring intersecting inequalities in human development. Research Report 04, Overseas Development Institute.

Lopus, S. and Frye, M. (2018). Visualizing Africa's Educational Gender Gap. *Socius: Sociological Research for a Dynamic World*, 4:1–2. doi: 10.1177/2378023118795956.

Mancini, L. (2008). Horizontal Inequality and Communal Violence: Evidence from Indonesian Districts. In Stewart, F., editor, *Horizontal Inequalities and Conflict*, pages 106–135. Palgrave Macmillan, London. doi: 10.1057/9780230582729_6.

Mancini, L., Stewart, F., and Brown, G. K. (2008). Approaches to the Measurement of Horizontal Inequalities. In Stewart, F., editor, *Horizontal Inequalities and Conflict*, pages 85–105. Palgrave Macmillan, London. doi: doi.org/10.1057/9780230582729_5.

Marrero, G. A. and Rodríguez, J. G. (2013). Inequality of opportunity and growth. *Journal of Development Economics*, 104:107–122. doi: 10.1016/j.jdeveco.2013.05.004.

McDaniel, A., DiPrete, T. A., Buchmann, C., and Shwed, U. (2011). The Black Gender Gap in Educational Attainment: Historical Trends and Racial Comparisons. *Demography*, 48(3):889–914. doi: 10.1007/s13524-011-0037-0.

McDoom, O. S., Reyes, C., Mina, C., and Asis, R. (2019). Inequality Between Whom? Patterns, Trends, and Implications of Horizontal Inequality in the Philippines. *Social Indicators Research*, 145(3):923–942. doi: 10.1007/s11205-018-1867-6.

Mittleman, J. (2021). Intersecting the Academic Gender Gap: The Education of Lesbian, Gay and Bisexual America. Technical report, SocArXiv.

Montalvo, J. G. and Reynal-Querol, M. (2003). Religious polarization and economic development. *Economics Letters*, 80(2):201–210. doi: 10.1016/S0165-1765(03)00080-6.

Montalvo, J. G. and Reynal-Querol, M. (2005). Ethnic polarization, potential conflict, and civil wars. *American Economic Review*, 95(3):796–816. doi: 10.1257/0002828054201468.

Muller, C. (2017). Ethnic Horizontal Inequity in Indonesia. Working paper.

Parker, P. D., Van Zanden, B., Marsh, H. W., Owen, K., Duineveld, J. J., and Noetel, M. (2020). The Intersection of Gender, Social Class, and Cultural Context: a Meta-Analysis. *Educational Psychology Review*, 32(1):197–228. doi: 10.1007/s10648-019-09493-1.

Piketty, T. and Saez, E. (2014). Inequality in the long run. *Science*, 344(6186):838–843. doi: 10.1126/science.1251936.

Ravallion, M. (2014). Income inequality in the developing world. *Science*, 344(6186):851–855. doi: 10.1126/science.1251875.

Riegle-Crumb, C. (2006). The Path through Math: Course Sequences and Academic Performance at the Intersection of Race-Ethnicity and Gender. *American Journal of Education*, 113(1):101–122. doi: 10.1086/506495.

Sen, G., Iyer, A., and Mukherjee, C. (2009). A Methodology to Analyse the Intersections of Social Inequalities in Health. *Journal of Human Development and Capabilities*, 10(3):397–415. doi: 10.1080/19452820903048894.

Shields, S. A. (2008). Gender: An Intersectionality Perspective. *Sex Roles*, 59(5):301–311. doi: 10.1007/s11199-008-9501-8.

Shorrocks, A. F. (1984). Inequality Decomposition by Population Subgroups. *Econometrica*, 52(6):1369. doi: 10.2307/1913511.

Stewart, F. (2009). Horizontal Inequality: Two Types of Trap. *Journal of Human Development and Capabilities*, 10(3):315–340. doi: 10.1080/19452820903041824.

Strid, S., Walby, S., and Armstrong, J. (2013). Intersectionality and multiple inequalities: Visibility in british policy on violence against women. *Social Politics*, 20(4):558–581. doi: 10.1093/sp/jxt019.

Tetteh-Baah, S. K. (2019). *Measurement and impact of horizontal inequality*. PhD thesis, ETH Zurich.

UNSDG (2022). Leave No One Behind. Retrieved from: https://unsdg.un.org/2030-agenda/universal-values/leave-no-one-behind.

Walby, S., Armstrong, J., and Strid, S. (2012). Intersectionality: Multiple Inequalities in Social Theory. *Sociology*, 46(2):224–240. doi: 10.1177/0038038511416164.

Tables and Figures

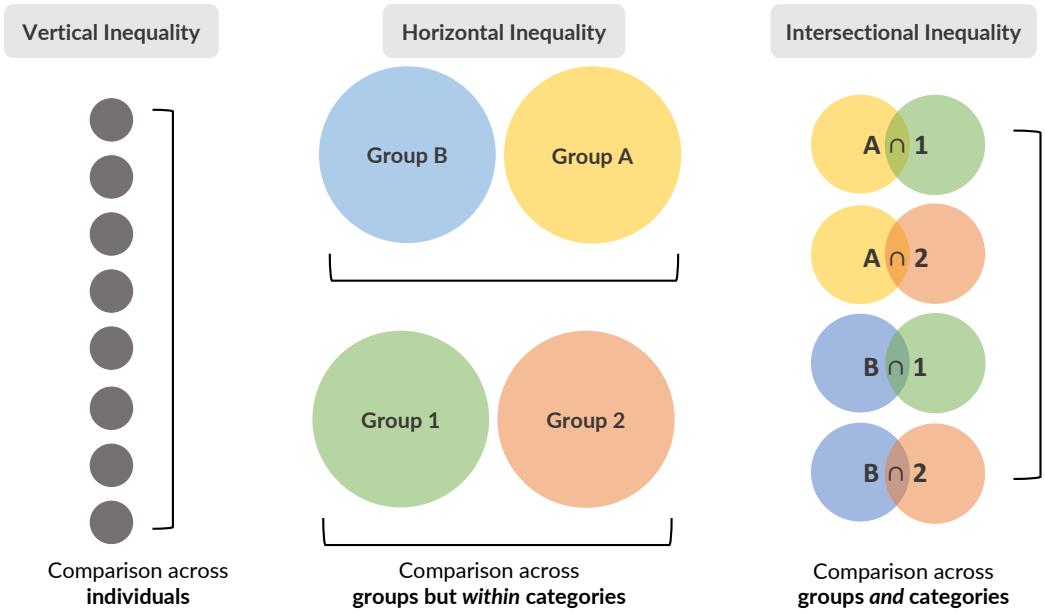

Figure 1: Concepts of inequality measurements

Notes: Authors' own representation adopted from Lenhardt and Samman (2015)

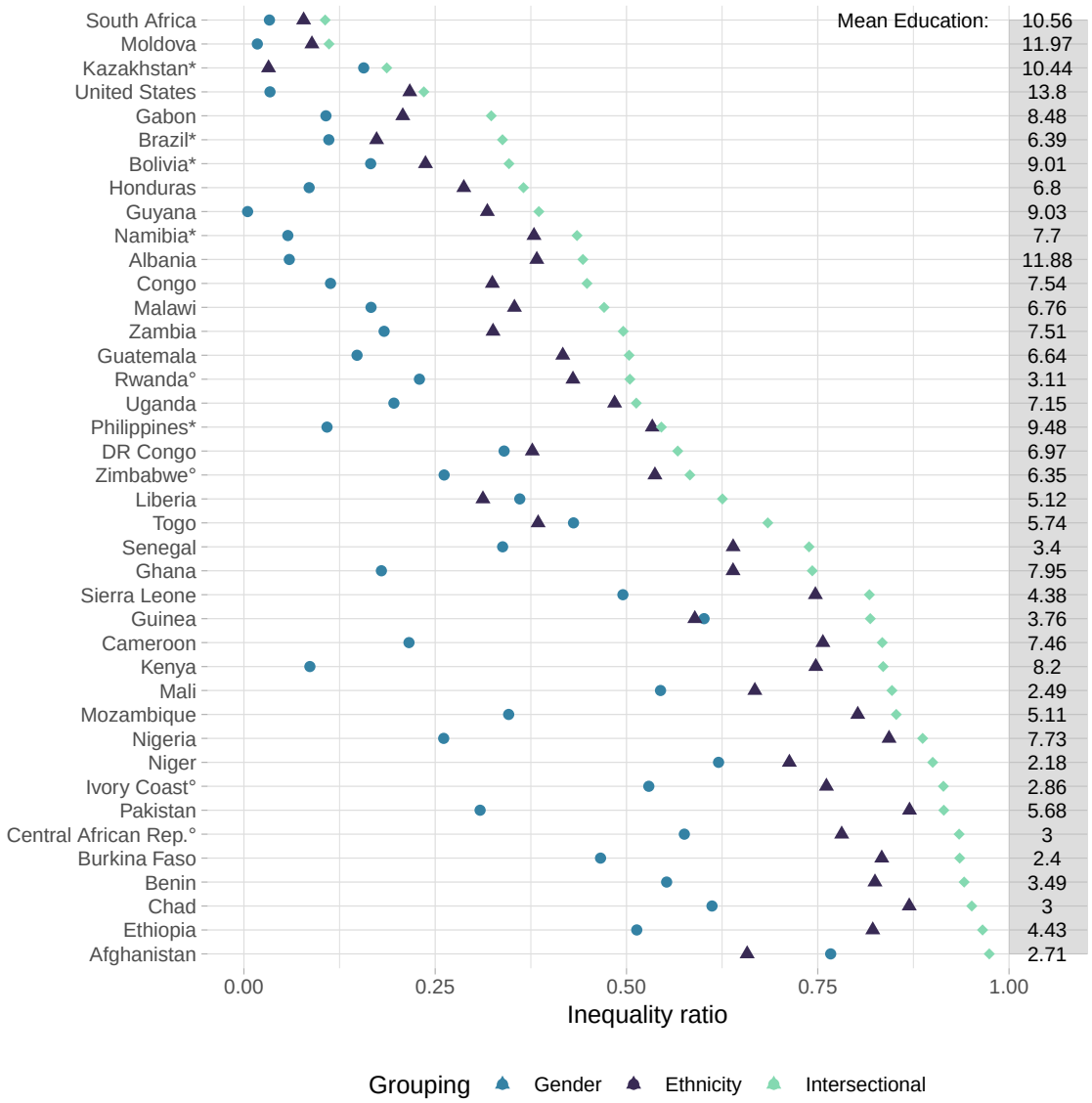

Figure 2: Inequality in education (years of schooling) by gender, ethnicity and intersecting groups

Notes: Aggregated data by country of n =2,689,279 individuals of the last available cohort; no mark means 1980 and younger cohort; countries marked with $*$ means 1970-1979 cohort; countries marked with °, the cohort born in 1969 and earlier was used. Using DHS sample weights, estimates show inequality ratios between groups with the lowest and highest average years of education. A value of zero implies parity and a value of one implies total inequality between the two most extreme groups. Sources: DHS 1992-2019 and US CPS 2019.

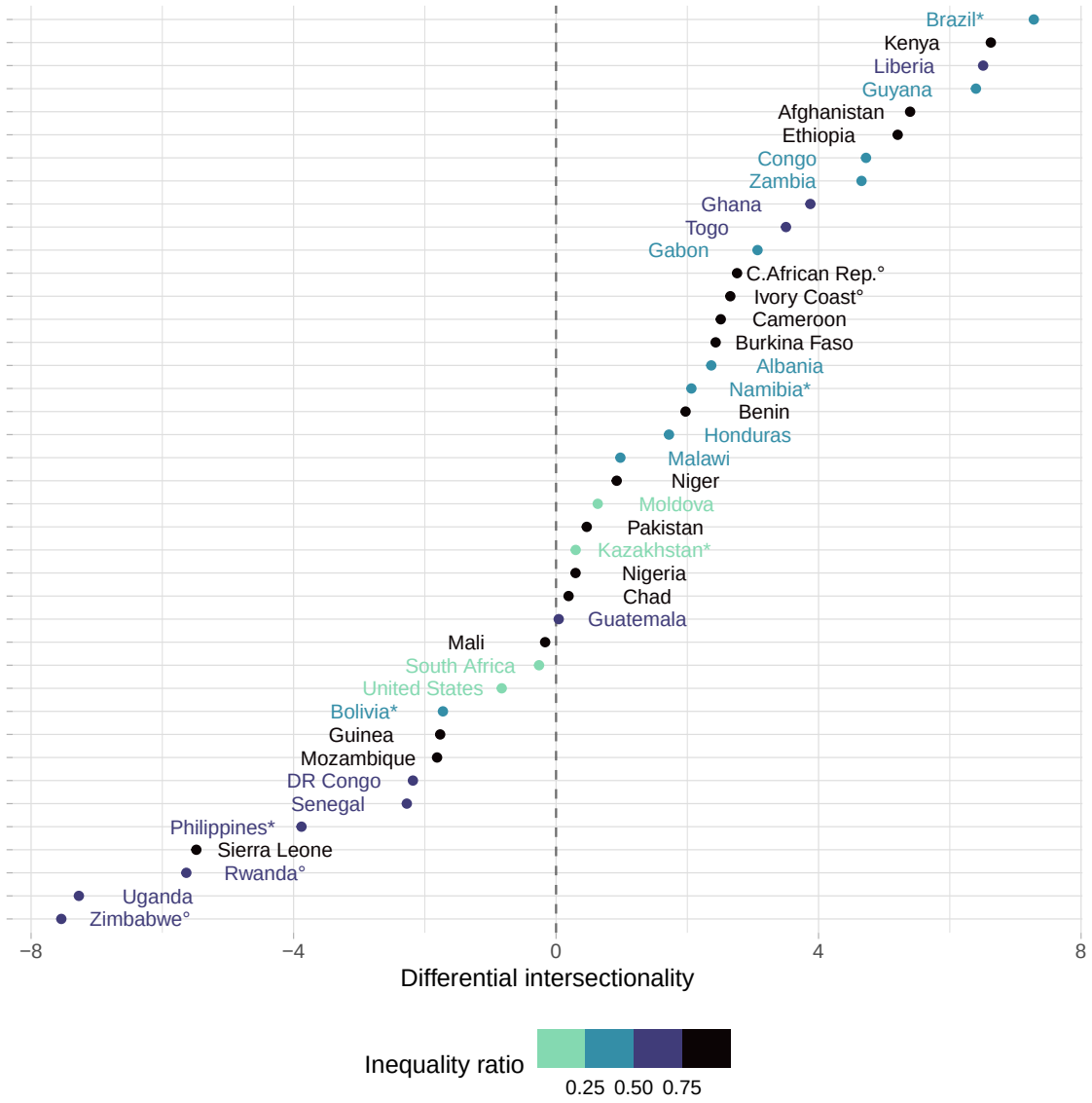

Figure 3: Differential intersectionality (difference between counterfactual and total intersectional inequality

Notes: Aggregated data by country of n =2'689'279 individuals of the last available cohort; no mark means 1980 and younger cohort; countries marked with $*$ means 1970-1979 cohort; countries marked with °, the cohort born in 1969 and earlier was used. Estimates show the relative difference (in percent) between counterfactual and total intersectional inequality ratios. Values above zero indicate higher total intersectional inequality (lower IR) than would be the case with constant relative gender gaps across ethnic groups. Sources: DHS 1992-2019 and US CPS 2019.

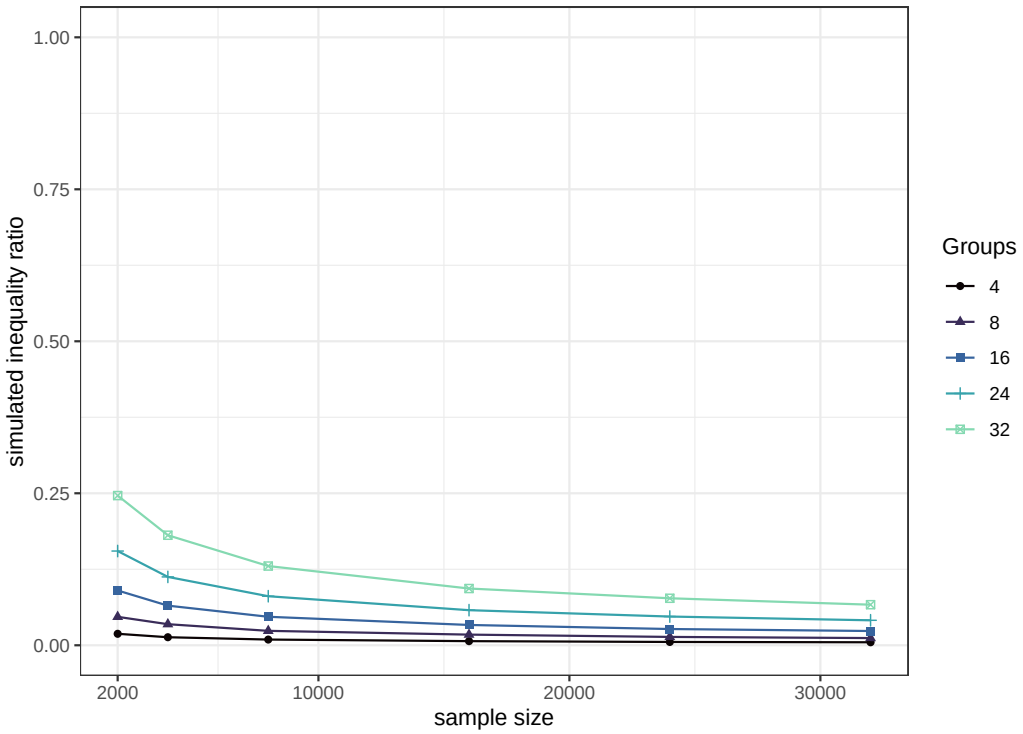

Figure 4: Simulation with varying group size and sample size

Notes: Based on averages across 1000 iterations. Education data is drawn randomly from a truncated normal distribution bounded between 0 and 17 years with $\mu = 5.5$ and $\sigma = 3$, corresponding to the average education and standard deviation across all countries.

Table 1: Descriptive statistics by birth cohort brackets

Characteristic	**-1969**, N = 30	**1970-1979**, N = 30	**1980-**, N = 30
Education (yrs)	5.0 (3.4)	5.6 (3.4)	6.4 (3.3)
IR(gender)	0.37 (0.21)	0.33 (0.21)	0.28 (0.20)
IR(ethnicity)	0.60 (0.25)	0.56 (0.25)	0.53 (0.25)
IR(gender*ethnicity)	0.73 (0.25)	0.68 (0.27)	0.64 (0.26)
Female (%)	0.62 (0.08)	0.72 (0.07)	0.73 (0.08)
No. of ethnic groups	8 (4)	8 (4)	8 (4)
Sample size	34,320 (133,352)	22,700 (67,330)	29,491 (104,609)
Group size low	940 (1,766)	1,019 (1,717)	1,468 (2,952)
Group size high	501 (579)	411 (472)	536 (1,037)

Notes: Statistics report the Mean (SD); Median (IQR) for no. of ethnic groups, aggregated by birth cohort brackets; n =2,689,279 individuals older than 25 years and younger than the birth cohort of 1920. For comparability, only countries with observations in all three cohort brackets are included. Sources: DHS 1992-2019 and US CPS 2019.

Table 2: Correlates of group inequality in education (OLS)

Outcome:	IR(gen:eth)		Counterfactual		Differential	
	(A)	(B)	(C)	(D)	(E)	(F)
Panel A: Education Variables						
Mean education (yrs)[a]	-0.06586***	-0.06057***	-0.06637***	-0.06071***	0.00051	0.00014
	(0.00609)	(0.00898)	(0.00595)	(0.00857)	(0.00107)	(0.00149)
Africa		0.08308		0.08917		-0.00609
		(0.06875)		(0.06591)		(0.01078)
Cohort 1970-1979[c]		-0.00590		-0.00230		-0.00360
		(0.01515)		(0.01227)		(0.00692)
Cohort 1980-[c]		0.01667		0.01850		-0.00183
		(0.02341)		(0.02147)		(0.00785)
(Intercept)	1.04552***	0.95336***	1.03457***	0.93423***	0.01094	0.01914
	(0.03217)	(0.08883)	(0.03067)	(0.08394)	(0.00712)	(0.01364)
Adj. R2	0.671	0.680	0.693	0.705	-0.007	-0.029
Panel B: Sample Size						
No. of ethnic groups	0.03251***	0.02387*	0.03242***	0.02303*	0.00008	0.00084
	(0.01155)	(0.01380)	(0.01132)	(0.01345)	(0.00112)	(0.00140)
Sample size[b]	-0.00066***	-0.00046***	-0.00062***	-0.00041**	-0.00004***	-0.00005***
	(0.00008)	(0.00017)	(0.00008)	(0.00016)	(0.00001)	(0.00002)
Africa		0.16692		0.17974		-0.01282
		(0.10911)		(0.10737)		(0.01074)
Cohort 1970-1979[c]		-0.05618**		-0.05140**		-0.00478
		(0.02196)		(0.02077)		(0.00632)
Cohort 1980-[c]		-0.06737**		-0.06504**		-0.00233
		(0.03027)		(0.02863)		(0.00701)
(Intercept)	0.42949***	0.41264***	0.41530***	0.39251***	0.01419	0.02013**
	(0.09333)	(0.08708)	(0.09256)	(0.08397)	(0.00855)	(0.00940)
Adj. R2	0.279	0.335	0.277	0.343	-0.008	-0.014
Panel C: Group Size						
Group size lowest	-0.00001	-0.00001	-0.00001	-0.00001	-0.00001***	-0.00001***
	(0.00002)	(0.00001)	(0.00002)	(0.00001)	(0.00000)	(0.00000)
Group size highest	-0.00006	0.00001	-0.00007	0.00001	0.00001*	0.00001
	(0.00004)	(0.00004)	(0.00004)	(0.00004)	(0.00000)	(0.00000)
Africa		0.28996***		0.29703***		-0.00706
		(0.10315)		(0.10047)		(0.00830)
Cohort 1970-1979[c]		-0.03158		-0.02834		-0.00323
		(0.02284)		(0.02041)		(0.00674)
Cohort 1980-[c]		-0.03945		-0.04024		0.00079
		(0.03413)		(0.03198)		(0.00729)
(Intercept)	0.70760***	0.49110***	0.69168***	0.46834***	0.01592***	0.02276**
	(0.04686)	(0.10587)	(0.04736)	(0.10255)	(0.00491)	(0.00891)
Adj. R2	0.034	0.234	0.029	0.243	0.065	0.047
Num.Obs	105	105	105	105	105	105
Mean of DV	0.6635	0.6635	0.6495	0.6495	0.0139	0.0139

Notes: Aggregated country-cohort bracket level data of 40 countries; n=2'689'279 individuals 25 years and older born after 1920; cluster-robust standard errors on the country-level in parentheses. Sign. codes: $^{*}p < 10\%$, $^{**}p < 5\%$, $^{***}p < 1\%$.

[a] Within-country-cohort weighted mean years of education with DHS sampling weights;

[b] Corresponds to country-cohort sample size in units of 1000;

[c] Ref: Cohort -1969.

Appendix A Supplementary Tables

Table A1: Education Inequality Ratios by Birth Cohort and Country for Gender and Ethnicity

Cohort	Education	IR(gender)	IR(eth)	IR(gen:eth)	IR(gen:eth)'	Differential	obs.
Afghanistan							
1970-1979	1.6 (3.62)	0.81	0.63	0.99	0.93	0.07	10207
1980-	2 (3.94)	0.77	0.66	0.97	0.92	0.05	17528
Albania							
-1969	11.37 (4.04)	0.06	0.21	0.27	0.26	0.01	10061
1970-1979	11.35 (4)	0.01	0.25	0.30	0.25	0.04	7103
1980-	12.14 (4.64)	0.06	0.38	0.44	0.42	0.02	6443
Benin							
-1969	2.33 (4.06)	0.54	0.93	0.97	0.97	0.01	17112
1970-1979	2.19 (3.7)	0.56	0.85	0.94	0.94	0.01	19084
1980-	2.86 (4.51)	0.55	0.82	0.94	0.92	0.02	17737
Bolivia							
-1969	6.78 (5.1)	0.17	0.33	0.43	0.45	-0.02	8613
1970-1979	8.57 (4.92)	0.17	0.24	0.35	0.36	-0.02	6097
Brazil							
-1969	6.13 (4.23)	0.08	0.25	0.32	0.31	0.01	9031
1970-1979	6.94 (3.88)	0.11	0.17	0.34	0.27	0.07	894
Burkina Faso							
-1969	0.92 (2.8)	0.50	0.76	0.90	0.88	0.02	18787
1970-1979	1.54 (3.52)	0.56	0.76	0.89	0.90	-0.01	11479
1980-	1.96 (3.78)	0.47	0.83	0.94	0.91	0.02	4883
Cameroon							
-1969	5.24 (4.54)	0.31	0.89	0.95	0.93	0.03	13493
1970-1979	6.27 (4.58)	0.25	0.79	0.91	0.84	0.06	13365

Table A1: Education Inequality Ratios by Birth Cohort and Country for Gender and Ethnicity *(continued)*

Cohort	Education	IR(gender)	IR(eth)	IR(gen:eth)	IR(gen:eth)'	Differential	obs.
1980-	7 (4.99)	0.22	0.76	0.83	0.81	0.03	12775
Central African Republic							
-1969	2.45 (3.64)	0.58	0.78	0.93	0.91	0.03	4616
Chad							
-1969	1.19 (2.84)	0.72	0.87	0.98	0.96	0.02	9039
1970-1979	1.61 (3.3)	0.70	0.83	0.96	0.95	0.01	7536
1980-	2.44 (3.97)	0.61	0.87	0.95	0.95	0.00	7733
Congo							
-1969	8.13 (4.38)	0.28	0.34	0.52	0.53	0.00	5256
1970-1979	8.2 (3.75)	0.19	0.19	0.38	0.35	0.03	7173
1980-	8.32 (3.85)	0.11	0.32	0.45	0.40	0.05	3930
DR Congo							
-1969	6.28 (4.59)	0.46	0.43	0.70	0.69	0.01	6636
1970-1979	6.55 (4.45)	0.35	0.39	0.56	0.61	-0.05	9335
1980-	6.59 (4.58)	0.34	0.38	0.57	0.59	-0.02	9355
Ethiopia							
-1969	1.48 (3.35)	0.68	0.90	NaN	0.97	NaN	17072
1970-1979	2.34 (3.92)	0.49	0.88	0.99	0.94	0.05	20654
1980-	3.42 (4.69)	0.51	0.82	0.97	0.91	0.05	18944
Gabon							
-1969	8.6 (3.8)	0.24	0.29	0.58	0.46	0.13	2079
1970-1979	9.03 (3.87)	0.20	0.33	0.49	0.46	0.03	2680
1980-	9.81 (3.82)	0.11	0.21	0.32	0.29	0.03	2506
Ghana							
-1969	6.14 (5.18)	0.34	0.84	0.89	0.90	0.00	11024
1970-1979	6.88 (5.03)	0.28	0.68	0.85	0.77	0.08	6505
1980-	8.27 (5.17)	0.18	0.64	0.74	0.70	0.04	5177

Table A1: Education Inequality Ratios by Birth Cohort and Country for Gender and Ethnicity *(continued)*

Cohort	Education	IR(gender)	IR(eth)	IR(gen:eth)	IR(gen:eth)'	Differential	obs.
Guatemala							
-1969	4.26 (4.7)	0.24	0.58	0.73	0.68	0.06	3783
1970-1979	4.92 (4.74)	0.20	0.51	0.63	0.61	0.02	7471
1980-	6.38 (4.9)	0.15	0.42	0.50	0.50	0.00	10798
Guinea							
-1969	1.77 (4.24)	0.62	0.62	0.84	0.86	-0.01	4734
1970-1979	1.64 (3.95)	0.72	0.58	0.88	0.88	0.00	4253
1980-	3.03 (5.29)	0.60	0.59	0.82	0.84	-0.02	5811
Guyana							
-1969	7.92 (3.39)	0.03	0.34	0.39	0.36	0.02	2065
1970-1979	8.64 (3.28)	0.05	0.35	0.38	0.38	0.00	2327
1980-	9.26 (3.09)	0.00	0.32	0.39	0.32	0.06	1118
Honduras							
-1969	5.66 (4.57)	0.11	0.39	0.54	0.46	0.08	4657
1970-1979	6.44 (4.3)	0.06	0.32	0.38	0.37	0.01	7036
1980-	7.51 (4.34)	0.09	0.29	0.37	0.35	0.02	6768
Ivory Coast							
-1969	2.4 (3.91)	0.53	0.76	0.91	0.89	0.03	6088
Kazakhstan							
-1969	10.77 (2.51)	0.15	0.04	0.20	0.18	0.02	3533
1970-1979	10.77 (2.24)	0.16	0.03	0.19	0.18	0.00	875
Kenya							
-1969	6.55 (4.3)	0.25	0.84	0.93	0.88	0.05	19245
1970-1979	8.16 (3.94)	0.13	0.80	0.88	0.83	0.05	18720
1980-	8.84 (3.99)	0.09	0.75	0.84	0.77	0.07	15797
Liberia							
-1969	4.39 (5.16)	0.65	0.45	0.82	0.80	0.02	1402

Table A1: Education Inequality Ratios by Birth Cohort and Country for Gender and Ethnicity *(continued)*

Cohort	Education	IR(gender)	IR(eth)	IR(gen:eth)	IR(gen:eth)'	Differential	obs.
1970-1979	4.55 (5.01)	0.49	0.43	0.71	0.71	0.00	3187
1980-	5.59 (5.12)	0.36	0.31	0.63	0.56	0.07	3776
Malawi							
-1969	3.63 (3.65)	0.40	0.66	0.84	0.80	0.04	15246
1970-1979	4.56 (3.96)	0.34	0.56	0.72	0.71	0.02	21043
1980-	6.32 (3.89)	0.17	0.35	0.47	0.46	0.01	17567
Mali							
-1969	1.29 (3.11)	0.53	0.66	0.88	0.84	0.03	21039
1970-1979	1.3 (3.2)	0.54	0.62	0.84	0.83	0.01	15553
1980-	2.01 (3.99)	0.54	0.67	0.85	0.85	0.00	12203
Moldova							
-1969	11.5 (2.33)	0.03	0.07	0.10	0.10	0.00	4121
1970-1979	11.54 (2.52)	0.01	0.03	0.04	0.04	0.00	2320
1980-	11.78 (2.74)	0.02	0.09	0.11	0.10	0.01	258
Mozambique							
-1969	2.21 (2.94)	0.53	0.83	0.89	0.92	-0.03	8073
1970-1979	3.09 (3.49)	0.38	0.82	0.87	0.89	-0.02	5241
1980-	3.96 (3.93)	0.35	0.80	0.85	0.87	-0.02	3993
Namibia							
-1969	6.57 (4.24)	0.02	0.46	0.53	0.46	0.07	3882
1970-1979	8.11 (3.76)	0.06	0.38	0.44	0.41	0.02	1868
Niger							
-1969	0.61 (2.15)	0.40	0.82	0.88	0.90	-0.01	13895
1970-1979	1.18 (2.83)	0.54	0.73	0.85	0.88	-0.03	4970
1980-	1.17 (2.91)	0.62	0.71	0.90	0.89	0.01	1142
Nigeria							
-1969	5.51 (5.68)	0.37	0.90	0.95	0.94	0.02	18869

Table A1: Education Inequality Ratios by Birth Cohort and Country for Gender and Ethnicity *(continued)*

Cohort	Education	IR(gender)	IR(eth)	IR(gen:eth)	IR(gen:eth)'	Differential	obs.
1970-1979	6.44 (5.69)	0.30	0.89	0.94	0.92	0.02	35890
1980-	7.16 (5.86)	0.26	0.84	0.89	0.88	0.00	47233
Pakistan							
-1969	3.09 (4.61)	0.56	0.87	0.99	0.94	0.05	2838
1970-1979	3.88 (4.95)	0.51	0.80	0.94	0.90	0.04	5525
1980-	4.49 (4.96)	0.31	0.87	0.91	0.91	0.00	5456
Philippines							
-1969	8.92 (4.21)	0.07	0.61	0.63	0.64	-0.01	7089
1970-1979	9.9 (4)	0.11	0.53	0.55	0.58	-0.04	4765
Rwanda							
-1969	2.53 (3.28)	0.23	0.43	0.50	0.56	-0.06	4364
Senegal							
-1969	2.39 (4.23)	0.41	0.56	0.77	0.74	0.02	23810
1970-1979	2.92 (4.35)	0.40	0.60	0.75	0.76	0.00	25146
1980-	3.59 (4.86)	0.34	0.64	0.74	0.76	-0.02	35847
Sierra Leone							
-1969	2.87 (4.88)	0.53	0.82	0.93	0.91	0.02	6013
1970-1979	2.44 (4.37)	0.51	0.86	0.91	0.93	-0.02	12515
1980-	3.87 (5.1)	0.50	0.75	0.82	0.87	-0.05	17503
South Africa							
-1969	8.78 (4.31)	0.08	0.26	0.33	0.32	0.01	1154
1970-1979	10.02 (3.58)	0.01	0.10	0.11	0.11	0.00	2537
1980-	10.85 (2.69)	0.03	0.08	0.11	0.11	0.00	4221
Togo							
-1969	3.34 (4.1)	0.58	0.63	0.87	0.84	0.02	7482
1970-1979	3.8 (4.11)	0.54	0.63	0.85	0.83	0.03	4954
1980-	5.38 (4.7)	0.43	0.38	0.68	0.65	0.04	4077

Table A1: Education Inequality Ratios by Birth Cohort and Country for Gender and Ethnicity *(continued)*

Cohort	Education	IR(gender)	IR(eth)	IR(gen:eth)	IR(gen:eth)'	Differential	obs.
Uganda							
-1969	4.17 (3.92)	0.42	0.63	0.77	0.79	-0.01	7065
1970-1979	5.11 (4.29)	0.31	0.53	0.65	0.68	-0.03	7187
1980-	6.91 (4.5)	0.20	0.48	0.51	0.59	-0.07	10809
United States							
-1969	13.23 (2.87)	0.01	0.28	0.28	0.29	0.00	739514
1970-1979	13.5 (2.91)	0.03	0.26	0.28	0.28	-0.01	376732
1980-	13.65 (2.64)	0.03	0.22	0.24	0.24	-0.01	580778
Zambia							
-1969	6.44 (4.05)	0.30	0.47	0.64	0.62	0.02	14043
1970-1979	6.77 (3.87)	0.22	0.38	0.48	0.52	-0.03	13487
1980-	7.4 (4.01)	0.18	0.33	0.50	0.45	0.05	10080
Zimbabwe							
-1969	6.07 (4.02)	0.26	0.54	0.58	0.66	-0.08	4506

Notes:

Education reports mean (sd); IR(G) reports inequality ratios between the group with the highest and lowest average education.

Table A2: Names of Groups with the Lowest and Highest Education by Birth Cohort

Cohort	Gender low	Gender high	Ethnicity low	Ethnicity high	Intersect. low	Intersect. high
Afghanistan						
1970-1979	F	M	nuristani	tajik	F:nuristani	M:tajik
1980-	F	M	nuristani	uzbek	F:nuristani	M:uzbek
Albania						
-1969	F	M	other	albanian	F:other	M:albanian
1970-1979	F	M	other	albanian	M:other	M:albanian
1980-	F	M	other	albanian	M:other	M:albanian
Benin						
-1969	F	M	Peulh	Yoruba	F:Peulh	M:Yoruba
1970-1979	F	M	Peulh	Other	F:Peulh	M:Fon
1980-	F	M	Peulh	Fon	F:Peulh	M:Adja
Bolivia						
-1969	F	M	quechua	none	F:quechua	M:none
1970-1979	F	M	quechua	none	F:quechua	M:aymara
Brazil						
-1969	M	F	black/mixed/o...	white	M:black/mixed...	F:white
1970-1979	M	F	black/mixed/o...	white	M:black/mixed...	M:white
Burkina Faso						
-1969	F	M	Gurma	Gurunsi	F:Gurma	M:Bobo
1970-1979	F	M	Fulfuldé/Peul	Gurunsi	F:Fulfuldé/Peul	M:Bobo
1980-	F	M	Gurma	Bobo	F:Gurma	M:Lobi
Cameroon						
-1969	F	M	Arab-choa/Peu...	Côtier/Ngoe/O...	F:Biu-Mandara	M:Côtier/Ngoe...
1970-1979	F	M	Arab-choa/Peu...	Beti/Bassa/Mbam	F:Biu-Mandara	M:Beti/Bassa/...
1980-	F	M	Arab-choa/Peu...	Côtier/Ngoe/O...	F:Arab-choa/P...	M:Beti/Bassa/...
Central African Republic						

Table A2: Names of Groups with the Lowest and Highest Education by Birth Cohort *(continued)*

Cohort	Gender low	Gender high	Ethnicity low	Ethnicity high	Intersect. low	Intersect. high
-1969	F	M	haoussa	yakoma-sango	F:haoussa	M:yakoma-sango
Chad						
-1969	F	M	gorane	sara (ngambay...	F:gorane	M:sara (ngamb...
1970-1979	F	M	kanembou / bo...	sara (ngambay...	F:kanembou / ...	M:sara (ngamb...
1980-	F	M	kanembou / bo...	sara (ngambay...	F:kanembou / ...	M:sara (ngamb...
Congo						
-1969	F	M	Other non-Con...	Mbohhi	F:Other non-C...	M:Mbohhi
1970-1979	F	M	Other Congolese	Mbohhi	F:Other non-C...	M:Mbeti
1980-	F	M	Other non-Con...	Mbohhi	F:Other non-C...	M:Mbeti
DR Congo						
-1969	F	M	uele lac albert	bakongo	F:ubangi and ...	M:bakongo
1970-1979	F	M	uele lac albert	bakongo	F:uele lac al...	M:bas-kasai a...
1980-	F	M	uele lac albert	bakongo	F:ubangi and ...	M:cuvette cen...
Ethiopia						
-1969	F	M	Affar	Welaita	F:Gumuz	M:Welaita
1970-1979	F	M	Affar	Guragie	F:Berta	M:Nuer
1980-	F	M	Affar	Nuer	F:Affar	M:Nuer
Gabon						
-1969	F	M	kota-kele	fang	F:kota-kele	M:fang
1970-1979	F	M	kota-kele	fang	F:kota-kele	M:fang
1980-	F	M	kota-kele	fang	F:kota-kele	M:fang
Ghana						
-1969	F	M	gruma	ga/dangme	F:gruma	M:ga/dangme
1970-1979	F	M	gruma	akan	F:gruma	M:akan
1980-	F	M	gruma	akan	F:gruma	M:akan
Guatemala						
-1969	F	M	maya/other	ladina/mestiza	F:maya/other	M:ladina/mestiza

Table A2: Names of Groups with the Lowest and Highest Education by Birth Cohort *(continued)*

Cohort	Gender low	Gender high	Ethnicity low	Ethnicity high	Intersect. low	Intersect. high
1970-1979	F	M	maya/other	ladina/mestiza	F:maya/other	M:ladina/mestiza
1980-	F	M	maya/other	ladina/mestiza	F:maya/other	M:ladina/mestiza
Guinea						
-1969	F	M	peulh	soussou	F:peulh	M:other
1970-1979	F	M	peulh	soussou	F:guerzé	M:soussou
1980-	F	M	peulh	other	F:peulh	M:other
Guyana						
-1969	M	F	amerindian	african	F:amerindian	F:african
1970-1979	M	F	amerindian	african	F:amerindian	F:african
1980-	M	F	amerindian	african	M:amerindian	F:african
Honduras						
-1969	M	F	None	Other	M:None	F:Other
1970-1979	M	F	None	Other	M:None	F:Other
1980-	M	F	None	Other	M:None	F:Other
Ivory Coast						
-1969	F	M	Burkina-Faso	Ivorian	F:Burkina-Faso	M:Ivorian
Kazakhstan						
-1969	M	F	Other	Russian	M:Other	F:Russian
1970-1979	M	F	Other	Russian	M:Russian	F:Russian
Kenya						
-1969	F	M	turkana	kikuyu	F:turkana	M:kikuyu
1970-1979	F	M	somali	kikuyu	F:somali	M:kisii
1980-	F	M	somali	kisii	F:somali	M:kisii
Liberia						
-1969	F	M	Kpelle	Grebo	F:Kpelle	M:Other Kru
1970-1979	F	M	Kpelle	Other Kru	F:Kpelle	M:Other Kru
1980-	F	M	Kpelle	Other	F:Kpelle	M:Grebo

Table A2: Names of Groups with the Lowest and Highest Education by Birth Cohort *(continued)*

Cohort	Gender low	Gender high	Ethnicity low	Ethnicity high	Intersect. low	Intersect. high
Malawi						
-1969	F	M	sena	tumbuka	F:sena	M:tumbuka
1970-1979	F	M	sena	tumbuka	F:sena	M:tumbuka
1980-	F	M	sena	tumbuka	F:sena	M:nkondhe
Mali						
-1969	F	M	dogon	malinke	F:dogon	M:malinke
1970-1979	F	M	tamacheck	malinke	F:dogon	M:malinke
1980-	F	M	tamacheck	bobo	F:tamacheck	M:bobo
Moldova						
-1969	M	F	moldovan	other	M:moldovan	F:other
1970-1979	M	F	moldovan	other	M:other	F:other
1980-	M	F	moldovan	other	F:moldovan	F:other
Mozambique						
-1969	F	M	Chewa	Portuguese	F:Sena	F:Portuguese
1970-1979	F	M	Chewa	Portuguese	F:Chewa	M:Portuguese
1980-	F	M	Chewa	Portuguese	F:Chewa	M:Portuguese
Namibia						
-1969	M	F	kavango langu...	afrikaans	F:kavango lan...	M:afrikaans
1970-1979	M	F	kavango langu...	afrikaans	F:kavango lan...	M:afrikaans
Niger						
-1969	F	M	Touareg/Touar...	Other	F:Touareg/Tou...	M:Other
1970-1979	F	M	Touareg/Touar...	Other	F:Touareg/Tou...	M:Djerma
1980-	F	M	Touareg/Touar...	Djerma	F:Kanouri	M:Djerma
Nigeria						
-1969	F	M	Fulani	Yoruba	F:Fulani	M:Ijaw/Izon
1970-1979	F	M	Fulani	Yoruba	F:Fulani	M:Ijaw/Izon
1980-	F	M	Fulani	Igbo	F:Fulani	M:Yoruba

Table A2: Names of Groups with the Lowest and Highest Education by Birth Cohort *(continued)*

Cohort	Gender low	Gender high	Ethnicity low	Ethnicity high	Intersect. low	Intersect. high
Pakistan						
-1969	F	M	Barauhi	Urdu	F:Balochi	M:Urdu
1970-1979	F	M	Barauhi	Urdu	F:Barauhi	M:Urdu
1980-	F	M	Barauhi	Urdu	F:Barauhi	M:Urdu
Philippines						
-1969	M	F	maguindanaon	tagalog	F:maguindanaon	F:tagalog
1970-1979	M	F	maguindanaon	tagalog	F:maguindanaon	F:tagalog
Rwanda						
-1969	F	M	hutu	tutsi/other	F:hutu	M:tutsi/other
Senegal						
-1969	F	M	Poular	Other	F:Poular	M:Diola
1970-1979	F	M	Poular	Diola	F:Poular	M:Diola
1980-	F	M	Poular	Diola	F:Poular	M:Diola
Sierra Leone						
-1969	F	M	Fullah	Creole	F:Kono	M:Creole
1970-1979	F	M	Other	Creole	F:Other	M:Creole
1980-	F	M	Other	Creole	F:Other	M:Creole
South Africa						
-1969	M	F	black/african	white/coloure...	M:black/african	M:white/colou...
1970-1979	M	F	black/african	white/coloure...	M:black/african	M:white/colou...
1980-	M	F	black/african	white/coloure...	M:black/african	F:white/colou...
Togo						
-1969	F	M	para-gourma/akan	akposso/akebou	F:para-gourma...	M:adja-ewe/mina
1970-1979	F	M	para-gourma/akan	akposso/akebou	F:para-gourma...	M:akposso/akebou
1980-	F	M	para-gourma/akan	kabye/tem	F:para-gourma...	M:ana-ife
Uganda						
-1969	F	M	Ruanda-Rundi	baganda	F:moru-madi	M:baganda

Table A2: Names of Groups with the Lowest and Highest Education by Birth Cohort *(continued)*

Cohort	Gender low	Gender high	Ethnicity low	Ethnicity high	Intersect. low	Intersect. high
1970-1979	F	M	moru-madi	baganda	F:moru-madi	M:langi
1980-	F	M	Ruanda-Rundi	baganda	F:alur-acholi	M:alur-acholi
United States						
-1969	M	F	Other race	Japanese	M:Other race	F:Japanese
1970-1979	M	F	Other race	Japanese	M:Other race	M:Japanese
1980-	M	F	Other race	Chinese	M:Other race	F:Chinese
Zambia						
-1969	F	M	mbunda	lozi	F:mbunda	M:namwanga
1970-1979	F	M	mbunda	namwanga	F:mbunda	M:namwanga
1980-	F	M	mbunda	lozi	F:mbunda	M:other
Zimbabwe						
-1969	F	M	black	white	F:black	F:white

Notes:

Table reports the names of the groups with the lowest and the highest average education that were used to calculate inequality ratios.

Table A3: Survey Years and Ethnic Groups

Country	Survey years	Ethnic groups	N groups
Afghanistan	2015	tajik, other, pashtun, hazara, uzbek, turkmen, nuristani	7
Albania	2009, 2018, 2017	albanian, other	2
Benin	1996, 2001, 2006, 2012, 2011, 2018, 2017	Other, Yoa/Lokpa, Bariba, Fon, Yoruba, Peulh, Betamaribe, Dendi, Adja	9
Bolivia	2004, 2003	quechua, none, aymara, other	4
Brazil	1996	black/mixed/other, white	2
Burkina Faso	1992, 1999, 2003, 2010	Mossi, Bobo, Gurunsi, Fulfuldé/Peul, Lobi, Other, Gurma, Senufo, Bissa, Dagara	10
Cameroon	1998, 2004, 2011, 2018, 2019	Arab-choa/Peulh/Haoussa/Kanuri, Biu-Mandara, Other, Bantoïde South-West, Bamilike/Bamoun, Adamaoua-Oubangui, Grassfields, Côtier/Ngoe/Oroko, Beti/Bassa/Mbam, Kako/Meka/Pygmé	10
Central African Republic	1994	banda, mandjia, ngbaka-bantou, other, yakoma-sango, gbaya, mboum, haoussa, sara	9
Chad	1996, 2004, 2014, 2015	arabic, sara (ngambaye/sara madjin-gaye/mbaye), other, ouadaï / maba / massalit / mimi, hadjarai, gorane, kanembou / bornou / boudouma	7
Congo	2005, 2011, 2012	Kongo, Other Kongo, Other Congolese, Balari, Teke, Mbohhi, Mbeti, Other non-Congolese, Ubangi, Sangha	10
DR Congo	2007, 2014, 2013	cuvette central, bas-kasai and kwilu-kwngo, ubangi and itimbiri, kasai, katanga, tanganika, bakongo, basele-komo, maniema, kivu, uele lac albert, other	8

Table A3: Survey Years and Ethnic Groups *(continued)*

Country	Survey years	Ethnic groups	N groups
Ethiopia	2000, 2005, 2011, 2016	Tigray, Amhara, Affar, Oromo, Guragie, Welaita, Somali, Sidama, Berta, Kefficho, Other, Gumuz, Agew, Gamo, Highlands, Hadiya, Ometo-Gimira/Basketo, Nuer	18
Gabon	2012	kota-kele, other, nzabi-duma, shira-punu/vili, fang, mbede-teke	6
Ghana	1993, 1998, 2008, 2014	akan, guan, ewe, ga/dangme, other, mole-dagbani, grussi, gruma	8
Guatemala	2015, 2014	ladina/mestiza, maya/other	2
Guinea	1999, 2018	peulh, malinké, soussou, kissi, other, guerzé	6
Guyana	2009	amerindian, mixed/other, indian, african	4
Honduras	2012, 2011	None, Other indigenous, Other, Maya chorti, Lenca	5
Ivory Coast	1994	Other nationality, Burkina-Faso, Ivorian	3
Kazakhstan	1999	Kazakh, Russian, Other	3
Kenya	1999, 1998, 2003, 2009, 2008, 2014	kikuyu, meru/embu, luhya, luo, kamba, taita/taveta, other, somali, mijikenda/swahili, kisii, kalenjin, maasai, turkana, oromo/gabbra/borana	14
Liberia	2013	Other Kru, Grebo, Kpelle, Other Mande, Other, Bassa	6
Malawi	2000, 2004, 2005, 2010, 2016, 2015	other, tumbuka, tonga, nkondhe, ngoni, chewa, yao, sena, lomwe, mang'anja	10
Mali	1996, 1995, 2001, 2006, 2012, 2013, 2018	peulh, bambara, sarkole/soninke/marka, sonrai, other, malinke, senoufo/minianka, dogon, bobo, tamacheck	10
Moldova	2005	other, moldovan	2
Mozambique	1997, 2011	Makhuwa, Other, Sena, Tswa/Rhonga, Portuguese, Lomwe, Chewa, Changana, Chopi/Tonga, Shona/Ndau	10
Namibia	2000	other, kavango languages, oshiwambo, afrikaans, damara/nama, herero	6
Niger	1992, 1998, 2006	Other, Djerma, Peulh, Haoussa, Kanouri, Touareg/Touareg Bella	6

Table A3: Survey Years and Ethnic Groups *(continued)*

Country	Survey years	Ethnic groups	N groups
Nigeria	2008, 2013, 2018	Hausa, Other, Igbo, Yoruba, Fulani, Kanuri/Beriberi, Igala, Ibibio/Efik/Anaang, Tiv, Ijaw/Izon, Ekoi, Urhobo et. al	12
Pakistan	2012, 2013	Pushto, Punjabi, Siraiki, Other, Urdu, Barauhi, Sindhi, Balochi	8
Philippines	2003	tagalog, ilocano, cebuano, other, waray, other bisaya, ilonggo, bicolano, kapampangan, maguindanaon	10
Rwanda	1992	hutu, tutsi/other	2
Senegal	1993, 1997, 2005, 2011, 2010, 2014, 2015, 2016, 2017, 2018, 2019	Wolof, Other, Poular, Serer, Mandingue, Diola, Soninke	7
Sierra Leone	2008, 2013, 2019	Mende, Temne, Other, Mandingo, Limba, Loko, Kono, Creole, Sherbro, Fullah	10
South Africa	2016	black/african, white/coloured/other	2
Togo	1998, 2013, 2014	adja-ewe/mina, kabye/tem, para-gourma/akan, other, ana-ife, akposso/akebou	6
Uganda	1995, 2011, 2016	langi, alur-acholi, moru-madi, other, banyoro, banyankore, chiga, baganda, Ruanda-Rundi, masaba-luhya, batoro, teso-turakana, basoga, bagisu, Other Nyoro-Ganda	15
United States	2019	Black, White, Two or more races, Other Asian, American Indian, Other race, Chinese, Japanese	8
Zambia	1996, 2002, 2001, 2007, 2013, 2014	lala-bisa, lunda, lozi, nsenga, chewa, mambwe-lungu, bemba, lenje-tonga, tumbuka, namwanga, luvale, ngoni, ushi, kaonde-nkoya, other, chokwe-luchazi, lamba, mbunda	18
Zimbabwe	1994	black, other, white	3

Appendix B Supplementary Figures

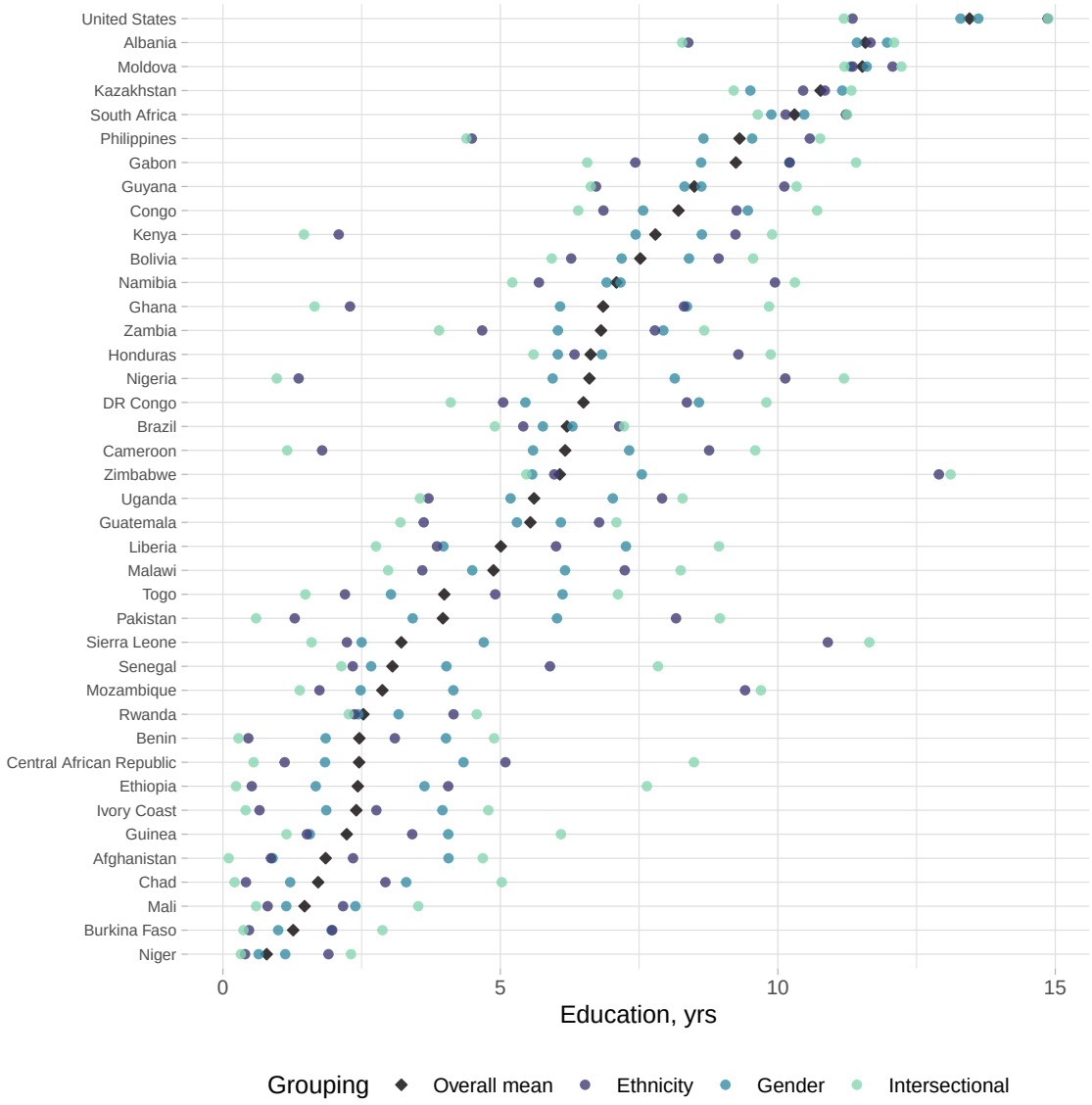

Figure B1: Average years in education by group

Notes: Pooled data from 40 countries; n =2,689,279 individuals older than 25 years and younger than the birth cohort of 1920. Estimates to the right of the overall mean represent the average education of the group with the highest education. Means are calculated using DHS sample weights. Sources: DHS 1992-2019 and US CPS 2019.

Appendix C Robustness Checks

Alternatively to the inequality ratio IR we can use a variant of the Theil Index, which is a special case of the General Entropy measures, as follows:

$$T(G) = \frac{1}{J}\sum_{j=1}^{J}\frac{s_j}{\mu_s} arsinh(\frac{s_j}{\mu_s}),$$

where s_j denotes the group averages, as previously defined for IR, J denotes the numbers of groups and μ the unweighted mean of all s_j. Instead of the natural logarithm, as is normally used for the Theil Index, we use an inverse hyperbolic sine transformation ($arsinh$). It has the advantage of being defined for zero values, which is a problem that we would inevitably run into with education data from low- and middle-income countries.

Repeating the regressions shown in Table 2, but with inequality measured by the Theil Index instead of the inequality ratio leads to the results reported in Table C1. The results remain fairly comparable.

Table C1: OLS regression of between-group Theil indices of education inequality

	Group Theil Index		
	(A)	(B)	(C)
Mean education (yrs)[a]	0.00559		
	(0.00708)		
Education Gini*b*	0.39624***		
	(0.11527)		
No. of ethnic groups		0.00296	
		(0.00439)	
Sample size[c]		-0.00011*	
		(0.00006)	
Group size lowest			-0.00001*
			(0.00000)
Group size highest			0.00000
			(0.00001)
Africa	-0.01366	0.03890	0.05458
	(0.01908)	(0.04250)	(0.03524)
Cohort 1970-1979[d]	-0.00474	-0.01576	-0.01250
	(0.00708)	(0.01020)	(0.01128)
Cohort 1980-[d]	-0.00910	-0.02475*	-0.01974
	(0.00905)	(0.01266)	(0.01356)
(Intercept)	0.76031***	0.94132***	0.95245***
	(0.09806)	(0.02742)	(0.03487)
Num.Obs	105	105	105
Adj. R2	0.629	0.061	0.055
Mean of DV	0.9762	0.9762	0.9762

Notes: Aggregated country-cohort bracket level data from 40 countries; n=2'689'279 individuals 25 years and older, and born after 1920; cluster-robust standard errors on the country-level in parentheses; Sign. codes: $^{*}p < 10\%$, $^{**}p < 5\%$, $^{***}p < 1\%$.

[a] Within-country-cohort weighted mean years of education with DHS sampling weights;

[b] Corresponds to country-cohort sample size in units of 1000;

[c] Ref: Cohort -1969.