Author's postprint:

PERMANYER, Iñaki; BOERTIEN, Diederik (2021) “Global Trends in Education Inequality: 1950-2010”. *Journal of Human Development and Capabilities*, 1-32 (Published Online: 10 April 2021). (ISSN: 1945-2829). DOI: https://doi.org/10.1080/19452829.2021.1911968

# Global Trends in Education Inequality: 1950-2010


## Abstract

*In this article, we document trends in global education inequality and its between and within-country components using absolute, intermediate and relative inequality measures. Existing studies have relied on converting categorical variables into continuous ones to estimate levels of inequality. Such procedures might not capture recent expansions in the duration of educational programs. We therefore compiled a database of 1164 datasets with information on years of education from a large set of countries across the world. According to our results, the absolute, intermediate and relative perspectives generate inconsistent narratives about recent trends in education inequality. While relative global inequality and its within- and between-country components have fallen monotonically during the last decades, absolute and intermediate inequality measures of actual years of education suggest that there might be a recent upsurge in global education inequality. Irrespective of the notion of inequality we adhere to, the bulk of global education inequality is explained more and more by variations within countries, with between-country inequality gradually losing ground as a source of variation. These findings suggest that the forces of globalization are rendering countries increasingly similar among them, and that whether global education inequality actually increments or decrements increasingly depends on what happens within those countries.*



Keywords: Education; Inequality; Global trends; Educational Attainment; Measurement;

## 1. Introduction

During several decades, education has been expanding all over the world. This expansion includes rising literacy rates (Crafts 2002) as well as increases in school enrollment rates and in completed years of primary, secondary and college education (Benavot and Riddle 1988; Meyer, Ramirez, and Soysal 1992; Barro and Lee 2001; Cohen and Soto 2007; Morrisson and Murtin 2009). Given the strong and beneficial effects of education on virtually all relevant life cycle events that social scientists are interested in, the spectacular expansion of education we have witnessed across the world's countries during the last decades has to be welcomed as a major social improvement.

While the levels and trends of average educational attainment indicators have been well documented, the study of educational *inequality* (i.e. how is education distributed across individuals) has not received much scholarly attention and is still quite poorly understood. Yet, the way in which education is distributed across individuals does have direct implications for their life chances. Increasing education inequality serves as a wellspring for increases in inequality in many other quality of life domains (higher variability in educational attainment is associated with higher dispersion in income and wages, higher job insecurity, and lower economic growth, occupational mobility, physical and mental health – see Dorius 2013, Ballarino et al 2014).

Empirical evidence on education inequality levels and trends paints a contentious and inconclusive picture, in large part due to the different countries, regions and time periods being compared, the inequality measures being used and the choice of basic unit of analysis. Previous studies investigating education inequality so far have mostly focused on educational differences among individuals *within* countries or on differences *between* countries, but surprisingly only a few of them have attempted to investigate *global*

education inequality (i.e. variations in individuals' education both within *and* between all world countries). In this paper, we document not only the joint evolution in within country, between country and global education inequality trends, but also investigate in detail the relationship between these trends and the expansion of primary, secondary and tertiary education. For that purpose, we develop and apply inequality decomposition techniques that allow going beyond purely descriptive results and analyze what factors are the most important drivers of education dispersion and its evolution over time. With these tools, we aim at investigating whether the expansion of successive education cycles has brought new layers of stratification, either locally or globally.

There are many reasons why one should be interested in studying *global* trends in education inequality. First, from an ethical perspective, if all human beings are entitled to equal opportunities, egalitarian principles should apply equally at the global and national level (e.g. Sen 2000, Bernstein 2011). Second, the study of global inequalities allows us to study global developments, such as the hotly debated consequences of economic globalization. In this regard, several studies have investigated recent trends in the global distribution of key well-being indicators. While some of these studies have investigated the global distribution of standards of living (see Goesling 2001, Bourguignon and Morrison 2002, Firebaugh and Goesling 2004, Anand and Segal 2015, Milanovic 2016) and length of life (Smits and Monden 2009, Edwards 2011), the empirical evidence on the global distribution of education is piecewise and not well articulated.

One of the key challenges that must be addressed in measuring education inequality is to generate reliable estimates of the distribution of education across individuals. Unfortunately, the variable 'years of schooling' – which is commonly used as a measure of educational attainment – is not available in many censuses and household surveys from which education information is usually collected. Instead, individuals are often grouped

into $k$ broad educational attainment categories (e.g. primary, secondary or tertiary education) thus limiting the applicability of standard inequality measures – which are based on cardinal variables. In order to overcome this limitation, the standard approach followed in the literature is to estimate education distributions imputing *average* years of schooling for each education cycle (e.g. Vinod et al 2001, Castelló and Doménech 2002, Benaabdelaali et al. 2012, Morrison and Murtin 2013). Surprisingly, no attempt has been made to assess whether such procedure (which eliminates the variation occurring within education categories) generates reliable estimates of true year of schooling distributions. In order to investigate this issue we have assembled a large dataset collecting information from different sources that allows studying for the first time *real* distributions of years of schooling around the globe. In this way, we will be able to assess whether the trends in education inequality based on the aforementioned imputation techniques cohere with the trends obtained from true years of schooling distributions. Looking at actual years of education provides important advantages as it allows capturing the possible increases in inequalities due to the increasing duration of tertiary education programs. In a recent Human Development Report, the UNDP (2019) expressed concerns that, even though greater global equality in basic capabilities has been achieved, inequality in enhanced capabilities is on the rise. Data on years of education will be particularly suited to test this argument.

Lastly, we critically discuss the implications of choosing different types of inequality measures – a seemingly technical issue traditionally overlooked in previous studies but that dramatically changes our findings. While most studies of education dispersion have relied on relative inequality indices, there is no theoretical nor practical reason to disregard their absolute (Kolm 1976a, 1976b) or intermediate (Krtscha 1994) counterparts. Indeed, there are several arguments to prefer absolute or intermediate

inequality measures when it comes to measure the variability of education distributions (Klasen et al 2018). Since the debate on whether inequality should be measured in relative, intermediate or absolute terms is contentious and unsettled, all our findings are reported using the three approaches. As is the case with the global distribution of income (Niño-Zarazúa, Roope and Tarp 2016), it turns out that the different perspectives lend support to inconsistent narratives about the recent trends in education inequality.

## 2. Background

What do we know about global trends in education inequality so far? Before addressing this question, we must clarify what we mean by 'the global distribution of education' and other inequality-related concepts, which, very often, can be confusing and lead to misunderstandings among researchers and other observers. Following analogous studies on the global distribution of income (e.g. Milanovic 2005) we define three concepts of inequality, distinguished by the population unit they refer to. The first one is the so-called 'Between-country education inequality' (BEI), which measures the inequality of a certain educational outcome across countries[1] (for instance: differences across countries in the average years of schooling). By definition, such measures compare country-level averages and disregard what happens with the distribution of education *within* countries. This is precisely the focus of our second concept of inequality, referred to as 'Within-country education inequality' (WEI), which measures the disparity in educational outcomes among citizens of a given country. Lastly, 'Global education inequality' (GEI) measures differences in educational levels across citizens all over the world. This concept

[1] In Dorius (2013), this type of inequality is referred to as 'International Education Inequality' (IEI).

of inequality – which is the key variable we want to analyze in this paper – takes into account variations occurring both between *and* within countries.

The vast majority of studies on education inequality have either focused on within country variations or on differences across countries (i.e. either on WEI or BEI separately). Many studies investigating education inequality are particularly interested in reporting WEI values across many countries and time periods, and in how the former relate to average educational levels (e.g. Vinod et al 2001; Castelló and Doménech 2002; Meschi and Scervini 2013; Ballarino et al 2014; Shukla & Mishra, 2019). Other studies have focused on between-country inequalities and in particular whether average education levels are converging or diverging over time across countries – an issue that has been thoroughly investigated for the international distribution of income (Barro and Sala-i-Martí 1992, Firebaugh 2003) and health (Firebaugh and Goesling, 2004, Wilson 2011).

Despite the importance of the third concept of inequality (which encompasses the previous two), it has only recently started to be analyzed in a handful of studies (see Benaabdelaali et al. 2012, Morrisson and Murtin 2013, Jordá and Alonso 2016). These studies either suggest an inverted U-shaped trajectory (Morrisson and Murtin 2013) or a negative relationship between inequality in imputed years of schooling and average imputed years of schooling (Benaabdelaali et al. 2012, Jordá and Alonso 2016). Two key difficulties have deterred the study of global education inequality. On the one hand, collecting roughly comparable educational attainment information across most world countries has never been an easy task – a problem that is gradually mitigated with the increasing availability of censuses, household surveys and international databases. On the other hand, the 'years of schooling' distributions often have to be estimated using

cardinalization techniques based on strong modeling assumptions[2]. We address these difficulties opting for a two-pronged strategy to maximize the use of existing data. First, we collect a new dataset on real years of schooling from a wide set of representative sources and use them to estimate trends in global education inequality.[3] Second, we cardinalize the abundant ordinal educational attainment data using standard imputation techniques to (i) investigate whether the resulting evidence coheres with the one generated from our dataset, and (ii) enlarge the geographic and temporal scope of our analysis (see the data and methods sections).

Due to the limitations of existing data and methods, existing empirical evidence of the *joint* evolution in WEI, BEI and GEI levels and its relationship with education expansion is fragmentary and non-articulated. In particular, the lack of data on years of education has prevented a proper investigation of whether new educational inequalities are replacing old ones (UNDP, 2019). There are several substantive reasons why new educational inequalities might keep emerging. Technological change could push demand for ever more specialized knowledge and skills (Goldin and Katz, 2008), and hence longer educational careers. Examples in this regard are Master and PhD programs. Another possibility is that when certain levels of educational attainment become universal, dominant social classes might push for new axes of stratification in educational systems

[2] A handful of studies have resorted to imputation and cardinalization techniques to estimate year of schooling distributions (see Benaabdelaali et al. 2012, Morrisson and Murtin 2013, Jordá and Alonso 2016). On the one hand, both Benaabdelaali et al. (2012) and Morrisson and Murtin (2013) combine educational attainment information with the average duration in each education cycle, thus assuming that there is no within-cycle variability and downwardly biasing their estimates. In addition, the last paper is restricted to 32 world macro regions – whose definitions are not clearly specified – thus eliminating a potentially important source of variation. On the other hand, Jordá and Alonso (2016) fit Generalized Gamma curves to educational attainment distributions to generate a continuous estimate of the variable 'years of schooling', but do not investigate whether the within-cycle heterogeneity they uncover is a realistic approximation of true years of schooling distributions.

[3] Meschi and Scervini (2014) present an international dataset on education inequality based on years of education, but it is heavily focused on European and OECD countries and it only includes information from one point in time.

in order to maintain their advantaged position (Alon, 2009; Lucas, 2001). These might include longer educational careers but can also emerge through horizontal forms of stratification including school type and field of study. Typical examples of such horizontal stratification are private schools and universities. One of the main aims of this paper is to get a better insight into the extent to which new educational inequalities are indeed emerging.

## 3. Methods

### *3.1 Conceptualization and Measurement of Educational Attainment*

Analyzing how educational attainment is distributed across individuals within and between countries is fraught with conceptual and practical difficulties. The first challenge one must face in this kind of study is the conceptualization and measurement of 'educational level'. The educational level indicator we will use in this paper is 'years of schooling' (henceforth YS). The advantage of using this variable is that it allows uncovering the entire education distribution with far more detail than broader educational attainment categories like 'primary, secondary or tertiary education completed'. On the minus side, the set of countries and years for which we can obtain the YS distribution is relatively small when compared to the countries and years for which we have ordinal education information. In order to circumvent this problem, researchers have used different imputation techniques to estimate YS distributions based on the percentages of population attaining primary, secondary or tertiary education. The standard approach consists in estimating the average years of schooling in each education cycle based on the

corresponding duration[4] – thus generating a discrete YS distribution (see Vinod et al 2001, Castelló and Doménech 2002, Morrison and Murtin 2013).

To compare the results of the data compiled on years of education to recent approaches that create imputed cardinal variables based on ordinal information (e.g. Castelló and Doménech 2002, Benaabdelaali et al. 2012) we also generate a simple discrete measure of years of education based on ordinal data provided by Barro and Lee.

*3.2 Inequality measurement*

When measuring the spread of YS distributions, there is often a normative debate on whether inequality should be measured in relative or absolute terms. The so-called relative inequality measures are invariant to any re-scaling of the distribution (i.e. inequality does not change when all outcomes are multiplied by the same constant). In contrast, absolute inequality measures are not mean-standardized and are invariant to translations (i.e. when the same constant is added to all elements of the distribution). Preferred by many economists for their ability to compare income distributions in different currencies, relative inequality measures are the workhorse of studies in global income inequality (Firebaugh and Goesling 2004, Anand and Segal 2015, Milanovic 2016). Yet, it is far from clear that relative measures should be preferred to absolute ones (e.g. in the context of this paper, it is not obvious that doubling the number of YS for everyone should leave inequality unaffected, because the differences between the most and the least educated would increase). Following Kolm (1976a, 1976b), in a situation where the variables we are dealing with are increasing (i.e., "growth"), absolute and

[4] Other authors, like Jordá and Alonso (2016), have fitted families of curves to discrete data points to generate continuous education distributions (see discussion section).

relative inequality indices are often referred to as 'leftist' and 'rightist' measures, respectively, and the choice among them is purely normative.

In an attempt to reconcile these extreme positions, different researchers suggested to introduce the so-called 'centrist' or 'intermediate' inequality measures, which increase (resp. decrease) whenever a distribution is scaled by a factor larger (resp. smaller) than one, and decrease (resp. increase) whenever the same quantity is added to (resp. subtracted from) all the elements of the distribution (Krtscha 1994, 2017). These measures strike a balance and try to find a fair compromise between the absolute and relative approaches by capturing the intuitions of both perspectives simultaneously. Intermediate inequality measures have been employed in the study of income inequality (e.g., see Subramanian (2015, 2017), Niño-Zarazúa et al (2016)), but we are not aware of their use in the study of education inequality. Given the lack of consensus regarding the invariance property an inequality index should satisfy,[5] in this paper we take a comprehensive perspective and present our findings using absolute/leftist, intermediate/centrist and relative/rightist inequality measures.

We have selected several inequality indices (two absolute, two relative and two intermediate ones) based on their popularity and their decomposability properties, which, as we show below, are very useful for the purposes of this paper. Assuming we have a distribution of years of schooling $\boldsymbol{y} = (y_1, \cdots, y_n)$ (where $y_i$ is the number of years of schooling of individual $i$), we start presenting the variance and the squared coefficient of variation. They are defined as

[5] For instance, some scholars suggest that inequality in non-monetary dimensions of well-being (like education or health) might be better captured with absolute or intermediate measures (Klasen et al 2018).

$$V(y_1, \cdots, y_n) = \frac{1}{n}\sum_{i=1}^{n}(y_i - \mu_y)^2 \qquad [1]$$

$$CV^2(y_1, \cdots, y_n) = \frac{V(y_1, \cdots, y_n)}{\mu_y^2} \qquad [2]$$

where $\mu_y$ represents the mean years of schooling of the distribution $\boldsymbol{y}$. Besides the aforementioned indices, in this paper we also take into consideration the absolute and relative versions of the Gini coefficient. They are defined as follows.

$$G_a(y_1, \cdots, y_n) = \frac{\sum_i \sum_j |y_i - y_j|}{2n^2} \qquad [3]$$

$$G_r(y_1, \cdots, y_n) = \frac{\sum_i \sum_j |y_i - y_j|}{2n^2\mu_y} \qquad [4]$$

The reasons for including the Gini coefficient in our analyses are: (i) it is a very popular measure of inequality whose relative version has already been used in well-known studies in the field (e.g. Vinod et al 2001, Castelló and Doménech 2002, Benaabdelaali et al. 2012), and (ii) it can be nicely decomposed according to the contribution of the different education cycles to total inequality (see below). We conclude this subsection introducing two well-known intermediate inequality indices: The Krtscha index (Krtscha 1994) and the intermediate Gini index (see Subramanian and Jayaraj 2013). They are defined as follows:

$$K(y_1, \cdots, y_n) = \frac{V(y_1, \cdots, y_n)}{\mu_y} \qquad [5]$$

$$G^*(y_1, \cdots, y_n) = G_a(y_1, \cdots, y_n) \cdot G_r(y_1, \cdots, y_n) \qquad [6]$$

Both indices are defined as the product of an absolute and a relative inequality measure: $K$ is the product between the standard deviation and the coefficient of variation and $G^*$ is

the product between the absolute and the relative Gini indices. In this way, both measures avoid the polar value judgements associated with the absolute and relative approaches. The following example illustrates how this intermediate position is arrived at. Starting from a given two-individual distribution, assume that an extra unit of the education variable we are dealing with becomes available. The $K$ index is defined in such a way that the extra unit to be distributed is split into two halves, one which preserves the existing pattern of the initial distribution (a relative notion of inequality) and the second which gives the same amount to the two individuals (an absolute notion of inequality). Following this approach, a compromise is found between the two extreme perspectives (see Subramanian 2015).

*3.3 Inequality decompositions*

The variance, the squared coefficient of variation and the $K$ index are well known for being additively decomposable.[6] In other words: when the population is partitioned in several mutually exclusive groups, the indices can be decomposed as:

$$V = V_W + V_B \qquad [7]$$

$$CV^2 = CV_W^2 + CV_B^2 \qquad [8]$$

$$K = K_W + K_B \qquad [9]$$

where $V_W, CV_W^2, K_W$ and $V_B, CV_B^2, K_B$ are the within and between groups inequality components, respectively. The within-group component is a weighted average of the inequality existing in each of the groups and the between-group component is obtained by calculating the inequality that would be observed in a hypothetical distribution where

[6] Other inequality indices are additively decomposable as well, like all the members of the family of Generalized Entropy measures (the Theil index being one of them). Yet, the substantive findings of the paper remain unaltered when switching to any of those measures.

each individual had the average years of schooling of the group to which s/he belonged (i.e. after eliminating within-group disparities). Using this exact decomposition, we can explore the extent to which global education inequality in years of schooling is driven by "location effects" (i.e. between-country inequality) or by "class effects" (within-country inequality).

While none of the Gini indices are additively decomposable, they can be broken down in another decomposition informing about the contribution of different education cycles to total inequality. Such decomposition is possible because the number of years that individual $i$ spent in school can be written as $y_i = p_i + s_i + t_i$, where $p_i, s_i, t_i$ are the years spent in primary, secondary and tertiary education, respectively (e.g. when someone has only attended primary school, $p_i > 0$ but $s_i = t_i = 0$). In Appendix 1, we show that the absolute and relative Gini coefficients can be rewritten as

$$G_r(\boldsymbol{y}) = \frac{\mu_p}{\mu_y}\overline{G_r}(\boldsymbol{p}) + \frac{\mu_s}{\mu_y}\overline{G_r}(\boldsymbol{s}) + \frac{\mu_t}{\mu_y}\overline{G_r}(\boldsymbol{t}) \qquad [10]$$

$$G_a(\boldsymbol{y}) = \overline{G_a}(\boldsymbol{p}) + \overline{G_a}(\boldsymbol{s}) + \overline{G_a}(\boldsymbol{t}) \qquad [11]$$

where $\mu_y$ is the overall mean of years of education, $\mu_p, \mu_s, \mu_t$ are the mean years of education spent in primary, secondary and tertiary education, and $\overline{G_a}(\boldsymbol{p}), \overline{G_a}(\boldsymbol{s}), \overline{G_a}(\boldsymbol{t})$, $\overline{G_r}(\boldsymbol{p}), \overline{G_r}(\boldsymbol{s}), \overline{G_r}(\boldsymbol{t})$ are the so-called 'absolute and relative pseudo-Gini coefficients' for the distributions of years of schooling spent in primary, secondary and tertiary education, respectively (see Appendix 1). These are the additive decomposition formulas that are used to assess the extent to which the three education cycles are contributing to global education inequality in years of schooling.[7]

[7] Interestingly, both decompositions give the same results for the absolute and relative cases (i.e. the contribution of primary, secondary and tertiary years of schooling to overall education inequality is the

In all our calculations, both BEI and GEI are weighted by the corresponding population sizes (both in the cardinal and ordinal settings)[8].

## 4. Data

The main contribution of this paper is to provide estimates of education inequality based on a new compilation of continuous data on years of education. A main empirical contribution of our study is that we collect data on years of education from various representative surveys as well as census data across the globe. Four cross-nationally comparable sources were identified that contain information on years of education: The IPUMS data, which is a harmonized database of mainly census data across the globe, the Demographic and Health Surveys (DHS), with survey data on primarily developing countries, the European Social Survey (ESS), and the International Social Survey Programme (ISSP). This compilation effort resulted in 1164 datasets, each referring to a given country-year.

We employ the complete compilation of 1164 datasets to examine the relationship between average levels of education and within-country levels of education inequality. Subsequently, we restrict the number of datasets to a consistent set of countries to compare estimates of between-country education inequality and global education inequality across two decades: 1994-2004 and 2005-2014. For 85 countries representative information on the distribution of years of education in each of the two periods is available (the selected countries are shown in Appendix 2). Datasets covering both periods were

same no matter if we use the relative or the absolute Gini coefficient). This is why the corresponding findings are not split into the absolute and relative cases.

[8] For some time, studies on global inequalities failed to weight by population size (this is the reason why some of the reported findings appear to be muddled – see Firebaugh (2003), Milanovic (2005), Dorius (2013)). Yet, failing to use population weights results in a strange concept of inequality where the importance of an individuals' experience is inversely proportional to the size of the country s/he happens to live in – so it has not been adopted in this paper.

required to come from the same data source (i.e. IPUMS, DHS, ESS, or ISSP) to safeguard comparability across time. If multiple datasets were available priority was given to IPUMS data followed by DHS, ESS, and ISSP data (given the generally larger sample sizes in the prioritized sources). If data was available for multiple years within each period, the datasets closest to the years 2000 and 2010 were chosen for each period respectively. The main variable of interest is years of education. In some cases education was top-coded in one period. In such cases, the variable was harmonized across periods by top-coding education in both periods.

In all of our analysis, we focus on respondents aged 30-34, as these are the youngest cohorts who have completed their educational careers. As such, zooming in on this age group allows us to detect the most recent changes in education inequality, and filters out the influence of age structures within countries on estimates of within-country education inequality. Across all 1164 datasets, final sample sizes of respondents aged 30-34 with information on years of education ranged from an average of 134 for data from the ISSP, 152 for data from the ESS, 2,067 for data from the DHS, to an average sample size of 118,374 for data from IPUMS.

For parts of the analysis, years of education were divided into stages (i.e. primary, secondary, and tertiary) based on the theoretical duration of each stage as reported by UNESCO's UIS.Stat database for the periods 1970-2016. Theoretical durations of educational stages were matched to each five-year birth cohort, and used to attribute years of education to each educational stage.[9]

[9] The following rules were employed to attribute years of education for each birth cohort. The age of cohorts is assumed to be the age of the third year of the age group. This is used to calculate birth year (e.g. for 15-19 in 2000, age 17, birth year=1983). Theoretical entry age in primary school for a given birth cohort: system that was in place at 4 years of age (e.g. for a person born in 1983 the entry age of the system reported in 1987 was used). Theoretical duration of primary education: system that was in place at age of entry. Theoretical duration of secondary education: system that was in place at (year-(age of entry into primary

*4.1 Ordinal Measures of Educational Attainment*

Even though our dataset of cardinal data on educational attainment is relatively well-suited to detect recent changes in education inequality, and hence, the possible emergence of new forms of education inequality, it has some shortcomings. The main limitation is the limited geographical and temporal scope of the data. We therefore complement our analysis with ordinal data, used in earlier studies, to monitor the possible influence of these limitations. Most previous studies have used Barro and Lee's data (2015), currently the most complete global database on educational attainment based on a compilation of censuses and surveys by UNESCO. The Barro and Lee data provide 7 ordinal categories of education for 146 countries during the period 1950-2010. We translated these categories into years of education using a uniform formula across countries (*No education:* 0 years; *Primary incomplete:* 3 years; *Primary complete:* 6, *Secondary incomplete: Secondary incomplete:* 9 years; *Secondary complete:* 12 years; *Tertiary incomplete:* 14 years; *Tertiary complete:* 16 years; robustness checks using alternative formulas, such as less years attributed to incomplete cycles, yielded very similar results). To be consistent with the previous dataset, we also focus here on individuals aged 30-34.

## 5. Empirical Findings

In this section, we present our empirical findings based on the systematic exploration of the dataset we assembled to estimate the YS distributions across countries and compare it to estimates based on ordinal education information reported in the Barro and Lee (2015) dataset. We start documenting the global education expansion and its regional

education + duration of primary education in years)). All educational system information for birth cohorts/relevant ages pre-1970s: 1970 values assigned.

trends during the period 1950-2010, for which the Barro and Lee data are suited best. Afterwards, we switch to our compilation of datasets to calculate the education inequality levels within countries, between countries and at a global scale. In each case, we present our estimates based on absolute and relative inequality measures. When feasible, we decompose inequality in its between and within-group components and assess the contribution of primary, secondary and tertiary education expansion to overall inequality levels.

*5.1 Education Expansion*

Figure 1 presents the share of the population having attended at least primary, secondary, and tertiary education for the world as a whole and in the seven regions in which it is partitioned (Appendix 3 shows the list of countries included in each region). Education has increased dramatically across all regions of the world. Primary education reached near universality, secondary education is no longer a minority phenomenon in all regions except Sub-Saharan Africa, and tertiary education has started to increase across the globe.

-Figure 1-

*5.2 Within-country Education Inequality (WEI)*

What can we say about the levels of education inequality within countries and their relationship with average educational attainment? Figure 2 shows the scatterplots comparing average years of schooling (horizontal axis) with the values of the variance (left panel), the $K$ index (central panel) and the squared coefficient of variation (right

panel) for the 1164 country-year samples we have assembled.[10] We can observe a strong negative relationship between years of schooling and *relative* education inequality (right panel): overall, the larger the average years of schooling, the more compressed the education distribution becomes. When the average years of schooling goes beyond 10, education inequality levels stabilize at very low levels. Tracking the evolution of countries with several observations over time we conclude that such general relationship applies to individual countries as well. These findings are very much in line with previous studies in the literature exclusively based on relative inequality measures (i.e. Castelló and Doménech 2002; Meschi and Scervini 2013).

-Figure 2-

Interestingly, we observe completely different patterns when measuring education variability with the variance (i.e. an absolute inequality index; see left panel in Figure 2). In that case, the relationship between average years of schooling and education inequality is much weaker. When the former ranges between 0 and 10, the cloud of points seem to follow an inverted U shape, but when it goes beyond 10 we observe no clearly discernible pattern. In general, the degree of variation across countries (i.e. the 'width of the cloud') is much larger than the one observed with the intermediate and, particularly, the relative inequality measures. Lastly, the trends for the intermediate inequality index $K$ are somewhere in between the two extreme cases. That is: we observe a strong negative relationship between average years of schooling and education inequality, which becomes weaker when the former go beyond the value of 10. Hence, even though relative

[10] For completeness, we have also looked at the relationship between average years of schooling and education inequality within countries using the absolute, intermediate and relative Gini coefficients. The results, which are shown in Figure A1 (see Appendix 4), are very similar to the ones shown in Figure 2.

inequality appears to mechanically decline with increasing levels of education, this does not imply that education becomes less dispersed.

Tracking the evolution of individual countries with repeated observations, we also observe an inverted U-shape relationship between average years of schooling below 10 and education inequality (i.e. in the initial and intermediate stages of education expansion an education Kuznets curve seems to emerge). Yet, what happens in the later stages of education expansion (i.e. average years of schooling above 10) is highly uncertain. The experience of individual countries is highly variable and no consistent nor statistically reliable pattern seems to emerge. Some countries like the United States still observe declines in the variance or the $K$ index as average years of schooling increase, but most countries tend to follow erratic patterns (e.g. Sweden, France, Canada, Spain). What happens with the education vanguard countries? The trajectory as observed for Finland (also observed for countries such as Denmark or Japan) might suggest that further education expansion could go in tandem with increases in education inequality. Yet, the reduced sample sizes and the small inconsistencies observed across data sources prevents reaching firm conclusions in that regard.

What about the contribution of primary, secondary and tertiary education to WEI levels? Figure 3 shows a ternary plot of the contribution of the three education cycles to the WEI levels for the 1164 country-year observation included in our cardinal dataset.[11] The different dots are shaded according to the average years of schooling. When the average years of schooling are small (light shade in the plot), primary education is the main contributor to inequality, followed by secondary education. As the average years of

[11] The position of each dot with respect to the vertices helps understanding how important is primary, secondary and tertiary education to explain observed inequality levels in years of schooling. The closer a dot is to a given vertex (say, P, S or T), the more important the corresponding education cycle is in explaining inequality. A hypothetical point in the middle of the triangle would represent a case where all education cycles contributed equally to observed inequality.

schooling increase, secondary education gradually becomes the most important driver of inequality. In later stages, further education expansion gives prominence to tertiary education as the main determinant of inequality, relegating primary education to the background. Put together, Figures 2 and 3 suggest that, even if education inequality in vanguard countries is mostly explained by the expansion of tertiary education, such expansion seems to be more equally distributed than the expansion of secondary education.

-Figure 3-

*5.3 Between-country Education Inequality (BEI)*

So far, we have explored education inequality levels among individuals belonging to the same country. What about differences among countries? Figure 4 plots the evolution of between country inequality for four indicators (overall mean years of schooling and mean years of schooling spent in primary, secondary and tertiary education) using leftist, centrist and rightist inequality measures (the Variance on the left panel, the $K$ index on the central panel, and the squared coefficient of variation on the right one)[12]. In each panel, we show the inequality levels derived from (a) the dataset we have assembled in this paper for the 2000–2010 period, and (b) from the imputations performed on the ordinal data from the BL dataset (spanning from 1950 to 2010).

Once again, the trends in BEI depend to a large extent on the choice of inequality measure. Using a relative measure like the squared coefficient of variation, we observe

[12] For completeness, we have also looked at trends in between-country education inequality corresponding to the absolute, intermediate and relative Gini coefficients. The results, which are shown in Figure A2 (see Appendix 4), are very similar to the ones shown in Figure 4.

convergence for the four BEI indicators as measured with the BL dataset, that is: differences across countries are decreasing substantially between 1950 and 2010 (see right panel). In addition, the average years of schooling spent in primary education are more equally distributed across countries than those spent in secondary education, which in turn are more equally distributed than those of tertiary education. Inspecting the 2000–2010 BEI trends derived from our dataset, we observe a highly consistent picture. The relative position and the downward slope of the curves is almost the same (except for the case of tertiary education), but the inequality estimates derived from our dataset are higher than the ones derived from BL data. Another important difference observed based on our data is that international education inequality in tertiary education appears to be increasing, whereas the BL estimates indicate decreases across educational levels.

Inspecting BEI trends with an absolute inequality index like the variance, we obtain completely different results (see left panel). Using the BL dataset, we observe that BEI levels are declining for primary education. For secondary education, they are increasing until the 1990s and decreasing thereafter. Between-country inequality is increasing for tertiary education across the entire observation period. The reported trends square well with previous findings using absolute inequality measures reporting the unequal diffusion of mass schooling across world countries (Dorius 2013). The relative position of the curves shown in the left panel of Figure 4 indicate that, so far, the diffusion of tertiary education has been more equitable than the diffusion of secondary education, which in turn has spread more equitably than primary education. If one adheres to the absolute notion of inequality, as education expands, the upper layers of education diffuse quicker across countries, with laggard countries catching up earlier than in the past. The 2000–2010 BEI trends we observe derived from our dataset are highly consistent with the ones derived from BL (the relative position and direction of the different curves is the same).

Once again, the BEI estimates derived from our dataset are higher than the ones derived from the BL dataset.

Lastly, the trends in intermediate inequality bear some similarities with the absolute and relative cases. On the one hand, when using the BL dataset, overall BEI and BEI in primary and secondary education follow generally declining trends – except for a small blip around the 1970s (in line with the results obtained for the Theil index). On the other hand, BEI in tertiary education tends to increase over time, a result that coheres with the variance trends. Inspecting the 2000-2010 results derived from our dataset, we observe trends that are consistent both with the absolute and relative cases (i.e. declining BEI for primary and secondary education and increasing BEI for tertiary education), although with very different levels of inequality.

-Figure 4-

*5.4 Global Education Inequality (GEI)*

In Figure 5 we plot the recent trends in global education inequality (and its within and between country components) based on both data sources. The left, central and right panels show the results for the variance, the $K$ index and the squared coefficient of variation, respectively.[13] When using a relative inequality measure like the squared coefficient of variation, we observe consistent patterns across data sources. In both cases, global educational inequality and its between and within country components decline over time. Throughout the observation period, education differences across countries are smaller than the differences within countries. In addition, the between-country component

[13] For completeness, we have also looked at the GEI trends corresponding to the absolute, intermediate and relative Gini coefficients. The results, which are shown in Figure A3 (see Appendix 4), are very similar to the ones shown in Figure 5.

becomes less relevant in determining global inequality trends over time (the contribution of the between country component hovers around 30% at the turn of the millennium, when it was around 50% back in the 1950s according to the BL dataset). Interestingly, our findings change substantially when using an absolute inequality measure like the variance. As can be seen in the left panel of Figure 5, GEI trends now differ substantially not only across data sources but also with respect to the findings based on the squared coefficient of variation. The estimates based on our database suggest that GEI has been on the rise, a trend that can be attributed to increases in within country education inequality. In contrast, the between-country component has remained stable over time, so its contribution to overall education inequality has become less important (from 44% around 2000 to 41% around 2010). On the other hand, the inequality estimates generated from the BL database follow a non-monotonic trend for the three types of inequality (WEI, BEI and GEI): increasing first, peaking, and finally declining. Despite such decline, as of 2010 the levels of WEI and GEI are higher than they were back in the 1950s. Lastly, the trends for the intermediate perspective (see central panel) are somewhere in between the previous two. Interestingly, the leftist, centrist and rightist perspectives coincide as regards the diminishing importance of between-country inequality in explaining global education inequality.

The lack of consistency between the sets of estimates generated from our 2000-2010 data and the BL data for the absolute and intermediate cases could indicate that the inequality levels derived from imputation techniques might be *under*-estimating true education inequality levels (a phenomenon that manifests as well in the Theil estimates shown in the right panel). A likely explanation is that ordinal measures of education do not capture recent expansions in the length of tertiary education.

-Figure 5-

Alternatively, the fact that the two data sources cover different sets of countries (see Appendices 2 and 3) might be an explanation for the discrepancy in GEI trends. To investigate whether this might be the case we have re-estimated the trends in GEI and its two subcomponents for the subset of countries included in both datasets simultaneously. The results, which are shown in Figure 6, confirm the findings reported in Figure 5. This strengthens the hypothesis that increases in more enhanced capabilities (UNDP, 2019) might lead to new increases in global education inequality after a period of sustained decreases in inequality.

-Figure 6-

What role have the different education stages played in explaining the trends in global education inequality? Figure 7 plots the contribution of the three education cycles to GEI levels using the same two data sources we employed before (based on a decomposition of the Gini index). As can be seen, the estimates are quite consistent across databases. In both cases, the contribution of primary education is declining while the contribution of tertiary education in increasing. The estimates obtained from the BL dataset provide a longer time perspective: between 1950 and 2010, the contribution of primary education to GEI inequality has decreased dramatically from 65% to 30%, it has been increasing and then plateauing around 55% for secondary education and has been increasing steadily for tertiary education – particularly after the 1970s – approaching 20% in 2010. This increasingly important role could strengthen and consolidate in the future as tertiary education expands and becomes responsible for an ever larger share of the global educational distribution.

-Figure 7-

## 6. Discussion and Concluding Remarks

Using the most up-to-date datasets on educational attainment, we have estimated global interpersonal education inequality levels from 1950 until 2010. For the first time, we document trends in global education inequality and its between and within-country components using absolute, intermediate and relative inequality measures. According to our results, relative global inequality and its two subcomponents, as captured by the squared coefficient of variation and the Gini coefficient, have fallen monotonically during the last decades (this is in line with the findings reported by Benaabdelaali et al 2012, Morrison and Murtin 2013, Jordá and Alonso 2016).

In contrast, absolute measures like the variance or the absolute Gini coefficient suggest that education inequality has followed an inverted U shape trend, and our compilation of data on actual years on education suggests that it might be currently increasing. Lastly, intermediate measures differ on their diagnosis regarding the trends in global education inequality: it declines according to the $K$ index and it follows an inverted U shape trend for the intermediate Gini coefficient. The conflicting evidence from absolute, intermediate and relative inequality measures highlights the importance of having an open discussion about the implications of normative notions of inequality when setting global development agendas, like United Nations' Sustainable Development Goals. Despite such discrepancies, the three perspectives agree on the increasing (resp. decreasing) importance of within-country (resp. between-country) inequality in explaining global trends education inequality. In this regard, our findings lend support to the hypothesis that the forces of globalization are rendering countries increasingly similar among them, and that whether global education inequality actually increments or decrements increasingly depends on what happens within countries.

The fact that education inequalities between countries are becoming increasingly small is not a priori surprising. Unlike income or wealth, education variables have natural ceilings

that mechanically reduce the corresponding levels of inequality as the mean of the distribution approaches those bounds. In a setting of education expansion, one should naturally expect that countries with lower educational attainment levels should eventually catch up with the more advanced ones, which can barely make any further improvements. At the same time, the analysis of this paper has shown that the extent to which between-country inequality has declined over time depends on the inequality measure chosen and whether data on years of education is used (instead of transformed categorical data on attainment). Our data on years of education has shown that between-country inequality might not have declined in absolute terms (i.e. using the Variance) over the last decades, as suggested by results based on transformed categorical data (See Figure 4).

Our empirical analyses indicate that recent education inequality trends crucially depend on the dataset that is used to estimate educational attainment distributions. One of the datasets we have used relies on imputing average years of schooling to the educational attainment categories available in the BL dataset (an approach that has also been used in Vinod et al 2001, Castelló and Doménech 2002, Benaabdelaali et al 2012). The other dataset is a multi-source collection of household surveys and censuses around the world containing information on *true* years of schooling distributions as reported by respondents (i.e. without relying on imputations). Part of the discrepancy in the 2000–2010 trends that obtains when using both sources of data (see the opposite directions in Figures 6 and 7) can be attributed to the depressing effect that imputations might have when measuring education inequality. When assuming that all those who attended a certain education cycle have stayed at school the same number of years, we are inevitably reducing variability in the years of schooling distributions.[14] In this regard, our findings suggest that imputation-

[14] Following this line of thought, Jordá and Alonso (2016) suggest fitting a family of generalized gamma density curves to discrete education distributions in order to uncover within cycle inequality. Since this approach tends to overestimate education inequality levels for those countries in which we have real years

based techniques like the ones that are routinely applied in the field might be underestimating true education inequality levels and potentially biasing the direction of the most recent trends. According to our results, imputation-based methods lead to a conclusion of decreasing global education inequality, whereas the imputation-free ones reach the opposite conclusion.

Summing up, while imputation-based techniques have been very useful to paint a faithful portrait of the broad-brush trends in education inequality from the 19th century to the turn of the millennium (Morrisson and Murtin 2013), our findings suggest they might be insufficient to give a more detailed account of the most recent trends in an increasingly interconnected and globalized world.

What do these findings tell us about whether tertiary and post-tertiary education generated new layers of inequality favoring certain sub-groups only – thus initiating a new upturn in global education inequality? The answers to these questions naturally depend on the basic unit of analysis (i.e. whether we compare countries, individuals within countries or individuals around the world) and on our normatively preferred notion of inequality (i.e. absolute, intermediate or relative). If we stick to the notion of relative inequality – the approach that economists have extensively employed to investigate *income* inequality – we inevitably reach the conclusion that education inequality in all its forms (within countries, between countries and globally) is declining. In contrast, the use of absolute and intermediate measures lends some support for the hypothesis that the expansion of tertiary education programs is leading to a new upsurge in education inequality. This conclusion applies to within-country inequalities and global education inequality,

of schooling distributions (results not shown here but available upon request), we have decided not to incorporate it in our analysis.

whereas trends in between-country inequality are estimated to be flat in recent years, even when using data on real years of education. Hence, new layers of inequalities are emerging within countries. As within-country inequality gains weight as a contributor to global inequality, these processes seem to have led to a recent upsurge in education inequality at the global level.

We conclude with a word of caution about the limits of our approach. Documenting how individuals' years of schooling are distributed across and within world countries is an important endeavor in and of itself. There is an extensive literature linking 'years of schooling' with all sorts of normatively desirable outcomes that social scientists and policy makers are interested in – including, but not limited to, job opportunities, earning potential, health outcomes and life satisfaction across the life cycle. Yet, educational attainment is not only about quantity, but also about quality. Very often, students do not acquire the skills they need by merely attending low quality schools with insufficient resources (Pritchett 2013). Unfortunately, currently existing databases only allow replacing 'years of education' with measures of education quality for a limited number of countries. Some sources, like the 'global data set on education quality', as harmonized by Altinok and colleagues (2018), allow estimating average student test scores across countries, but tell us nothing about within-country variability. Only a subset of the countries studied in this paper is part of that database, but for recent years a consistent subset of 69 countries with data on average test score outcomes from both 2000 and 2010 can be constructed. Between 2000 and 2010, absolute, intermediate and relative measures of between-country inequality in education quality declined, a result that coheres with the trends we have identified in this paper for 'years of education', but that ignores the heterogeneity that might exist within countries. In future research, a more thorough investigation will be required to complement these findings and investigate how

education quality is distributed both across and within countries in a truly global perspective.

## References


Alon, S. (2009). The evolution of class inequality in higher education: Competition, exclusion, and adaptation. American Sociological Review, 74(5), 731-755.

Altinok, N., Angrist, N. and Patrinos, HA. (2018). Global data set on education quality (1965-2015). Policy Research Working Paper 8314, World Bank.

Anand, S., and Segal, P. (2015). The global distribution of income. In Atkinson, A. and Bourguignon, F. (Eds.) Handbook of Income Distribution, Volume 2A (Chapter 11), North Holland.

Ballarino, G., Bratti, M., Filippin, A., Fiorio, C., Leonardi, M. and Scervini, F. (2014). Increasing educational inequalities?, in Salverda, W. et al. (Eds). Changing inequalities in rich countries: Analytical and comparative perspectives, Chapter 5. Oxford University Press, Oxford.

Barro, R. J., and Lee, J. W. (2001). International data on educational attainment: updates and implications. Oxford Economic papers, 53(3), 541-563.

Barro, R. J., and Lee, J. W. (2015). Education matters: global schooling gains from the 19th to the 21st century. Oxford: Oxford University Press.

Barro, R. J., and Sala-i-Martí, X. (1992). Convergence. Journal of Political Economy 100(2): 223-251.

Benaabdelaali, W., Hanchane, S., and Kamal, A. (2012). Educational Inequality in the World, 1950–2010: Estimates from a New Dataset. *Research on Economic Inequality*, 20) , 337-366.

Benavot, A., and Riddle, P. (1988). The expansion of primary education, 1870-1940: Trends and Issues. Sociology of Education, 61, 191-210.

Bernstein, A. (2011). Political cosmopolitanism. In: Chatterjee, D. (Ed.), Encyclopedia of Global Justice, Springer, London.

Bourguignon, F. and Morrisson, C. (2002). Inequality among world citizens: 1820-1992. American Economic Review, 92(4):727-744.

Castelló, A., and Doménech, R. (2002). Human capital inequality and economic growth: some new evidence. The Economic Journal, 112(478), C187-C200.

Cohen, D., and Soto, M. (2007). Growth and human capital: good data, good results. Journal of Economic Growth, 12(1), 51-76.

Crafts, N. (2002). The human development index, 1870–1999: Some revised estimates. European Review of Economic History, 6(3), 395-405.

Dorius, S. F. (2013). The Rise and Fall of Worldwide Education Inequality from 1870 to 2010 Measurement and Trends. Sociology of Education, 86(2), 158-173.

Edwards, R. D. (2011). Changes in world inequality in length of life: 1970–2000. Population and Development Review, 37(3), 499-528.

Firebaugh, G. (2003). The new geography of global income inequality. Harvard University Press.

Firebaugh, G. and Goesling, B. (2004). Accounting for the recent decline in global income inequality. The American Journal of Sociology, 110(2):283-312.

Goldin, C. and Katz, L. (2008). The race between education and technology. The Belknap Press of Harvard University Press.

Jordá, V. and Alonso, J. (2016). New estimates on educational attainment using a continuous approach (1970-2010). World Development, 90, 281-293.

Stephan Klasen, Giovanni Andrea Cornia, Rebeca Grynspan, Luis F. López- Calva, Nora Lustig, Augustin Fosu, Flora Myamba, Sripad Motiram, Andreas Peichl, Eldar Shafir, Ana Sojo, Ingrid Woolard (2018) "Economic Inequality and Social Progress," in Rethinking Society for the 21st Century: Report of the International Panel on Social Progress, International Panel on Social Progress, Cambridge: Cambridge University Press.

Kolm, S. Ch. (1976a). Unequal Inequalities I. Journal of Economic Theory, 12(3): 416-454.

Kolm, S. Ch. (1976b). Unequal Inequalities II. Journal of Economic Theory, 13(1): 82-111.

Krtscha, M. (1994). 'A New Compromise Measure of Inequality', in W. Eichhorn (ed.): Models and Measurement of Welfare and Inequality. Heidelberg: Springer-Verlag.

Krtscha, M. (2017). 'Some Axiomatics of Inequality Measurement, With Specific Reference to Intermediate Indices', Working Papers 445, ECINEQ, Society for the Study of Economic Inequality.

Lerman, R. and Yitzhaki, S. (1985). Income inequality effects by income sources: a new approach and applications to the United States. The Review of Economics and Statistics, 67 (1), 151-156.

Lucas, S. R. (2001). Effectively maintained inequality: Education transitions, track mobility, and social background effects. American journal of sociology, 106(6), 1642-1690.

Meschi, E., and Scervini, F. (2013). Expansion of schooling and educational inequality in Europe: the educational Kuznets curve revisited. Oxford Economic Papers, 66, 660-680.

Meschi, E., and Scervini, F. (2014). A new dataset on educational inequality. Empirical Economics, 47(2), 695-716.

Meyer, J. W., Ramirez, F. O., and Soysal, Y. N. (1992). World expansion of mass education, 1870-1980. Sociology of Education, 128-149.

Milanovic, B. (2005). Worlds apart: Measuring international and global inequality. Princeton University Press.

Milanovic, B. (2016). Global inequality. A new approach for the age of globalization. Harvard University Press.

Morrisson, C., and Murtin, F. (2009). The century of education. Journal of Human Capital, 3(1), 1-42.

Morrisson, C., and Murtin, F. (2013). The Kuznets curve of human capital inequality: 1870–2010. The Journal of Economic Inequality, 11(3), 283-301.

Niño-Zarazúa, M., Roope, L. and Tarp, F. (2016). Global inequality: Relatively lower, absolutely higher. Review of Income and Wealth (DOI: 10.1111/roiw.12240).

Pritchett, L. (2013). The rebirth of education. Schooling ain't learning. Center for Global Development, Washington DC.

Sen, A. (2000). Social justice and the distribution of income. In: Atkinson, A. & Bourguignon, F. (Eds.), Handbook of income distribution, Volume 1, Elsevier, Amsterdam.

Shorrocks, A. (1982). Inequality decomposition by factor components. Econometrica, 50(1), 193-211.

Shukla, V., & Mishra, U. S. (2019). Educational Expansion and Schooling Inequality: Testing Educational Kuznets Curve for India. Social Indicators Research, 141, 1265-1283.

Smits J. and Monden C. (2009). Length of life inequality around the globe. Social Science & Medicine, 68:1114-1123.

Subramanian, S. (2015). Assessing Inequality in the Presence of Growth. Review of Black Political Economy, 42(3): 179-199.

Subramanian, S. (2017). On Comprehensively Intermediate Measures of Inequality and Poverty, With an Illustrative Application to Global Data. Journal of Globalization and Development, 8(2): 1-18.

Subramanian, S. and D. Jayaraj (2013). The Evolution of Consumption and Wealth Inequality: A Quantitative Assessment. Journal of Globalization and Development, 4:253–281.

UNDP (2019). Beyond Income, Beyond Averages, Beyond Today: Inequalities in Human Development in the 21st Century. *Human Development Report 2019.* http://hdr.undp.org/sites/default/files/hdr2019.pdf

Vinod, T., Yan, W. and Xibo, F. (2001). Measuring education inequality. Gini coefficients of education. Policy Research Working Paper 2525. The World Bank.

Wilson, C. (2011). Understanding global demographic convergence since 1950. Population and Development Review 37(2):375-388.

## Appendices

*Appendix 1. Calculating the contribution of education cycles to overall education inequality*

The number of years that individual $i$ has spent in school can be written as

$$y_i = p_i + s_i + t_i \qquad [A1]$$

where $p_i, s_i, t_i$ are the years spent in primary, secondary and tertiary education, respectively. The distribution of years of schooling spent in primary, secondary, tertiary and overall education will be denoted as $\boldsymbol{p}, \boldsymbol{s}, \boldsymbol{t}$ and $\boldsymbol{y}$, respectively. Following Shorrocks (1982:195), if the years of schooling distribution is ordered so that$y_1 \leq y_2 \leq \cdots \leq y_n$, then the corresponding relative Gini coefficient can be written as

$$G_r(\boldsymbol{y}) = \frac{2}{n^2\mu_y}\sum_{i=1}^{n}\left(i - \frac{n+1}{2}\right)y_i \qquad [A2]$$

where $n$ is the number of individuals and $\mu_y$ is the mean of the years of schooling distribution. Plugging equation [A1] into equation [A2] we obtain

$$\begin{aligned} G_r(\boldsymbol{y}) &= \frac{2}{n^2\mu_y}\sum_{i=1}^{n}\left(i - \frac{n+1}{2}\right)(p_i + s_i + t_i) \\ &= \frac{\mu_p}{\mu_y}\overline{G_r}(\boldsymbol{p}) + \frac{\mu_s}{\mu_y}\overline{G_r}(\boldsymbol{s}) + \frac{\mu_t}{\mu_y}\overline{G_r}(\boldsymbol{t}) \qquad [A3] \end{aligned}$$

where $\mu_p, \mu_s$ and $\mu_t$ are the means of the $\boldsymbol{p}, \boldsymbol{s}$ and $\boldsymbol{t}$ distributions, and

$$\begin{cases} \overline{G_r}(\boldsymbol{p}) = \dfrac{2}{n^2\mu_p}\displaystyle\sum_{i=1}^{n}\left(i - \frac{n+1}{2}\right)p_i \\ \overline{G_r}(\boldsymbol{s}) = \dfrac{2}{n^2\mu_s}\displaystyle\sum_{i=1}^{n}\left(i - \frac{n+1}{2}\right)s_i \qquad [A4] \\ \overline{G_r}(\boldsymbol{t}) = \dfrac{2}{n^2\mu_t}\displaystyle\sum_{i=1}^{n}\left(i - \frac{n+1}{2}\right)t_i \end{cases}$$

are the so-called pseudo-Ginis for distributions $\boldsymbol{p}, \boldsymbol{s}$ and $\boldsymbol{t}$ respectively (see Shorrocks 1982:196 and Lerman and Yitzhaki 1985:152). These pseudo-Ginis are *not* the conventional Gini values $G_r(\boldsymbol{p}), G_r(\boldsymbol{s}), G_r(\boldsymbol{t})$, since the weights attached to the corresponding $p_i, s_i$ and $t_i$ in equation [A4] correspond to the rank of individual '$i$' in the distribution $\boldsymbol{y}$, which in general is not the same as its rank in the distributions $\boldsymbol{p}, \boldsymbol{s}$ and $\boldsymbol{t}$. Equation [A3] (which coincides with equation [10] in the main text) provides a natural additive decomposition rule for the relative Gini index where the contributions of the different sources are clearly established. In order to derive the additive decomposition of the absolute Gini coefficient $G_a$ shown in equation [11] we follow the same procedure as before (the only difference being that the mean distributions are not appearing in the denominators).

*Appendix 2: Datasets selected:*

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Argentina | 2001 | IPUMS | Cote Ivoire | 1998 | DHS | Haiti | 2000 | DHS | Mali | 1998 | IPUMS | Rwanda | 2000 | DHS |
| Argentina | 2010 | IPUMS | Cote Ivoire | 2011 | DHS | Haiti | 2012 | DHS | Mali | 2009 | IPUMS | Rwanda | 2010 | DHS |
| Armenia | 2000 | DHS | Cyprus | 1999 | ISSP | Hungary | 2002 | ESS | Mexico | 2000 | IPUMS | Slovakia | 2004 | ESS |
| Armenia | 2010 | DHS | Cyprus | 2009 | ISSP | Hungary | 2010 | ESS | Mexico | 2010 | IPUMS | Slovakia | 2010 | ESS |
| Australia | 1999 | ISSP | Czech Republic | 2002 | ESS | Iceland | 2004 | ESS | Mozambique | 1997 | DHS | Slovenia | 2002 | ESS |
| Australia | 2009 | ISSP | Czech Republic | 2010 | ESS | Iceland | 2012 | ESS | Mozambique | 2009 | DHS | Slovenia | 2010 | ESS |
| Austria | 2002 | ESS | Denmark | 2002 | ESS | India | 1998 | DHS | Namibia | 2000 | DHS | South Africa | 2001 | IPUMS |
| Austria | 2006 | ESS | Denmark | 2010 | ESS | India | 2005 | DHS | Namibia | 2013 | DHS | South Africa | 2011 | IPUMS |
| Bangladesh | 2001 | IPUMS | Dominican Rep. | 2002 | IPUMS | Indonesia | 2002 | DHS | Nepal | 2001 | DHS | Spain | 2002 | ESS |
| Bangladesh | 2011 | IPUMS | Dominican Rep. | 2010 | IPUMS | Indonesia | 2012 | DHS | Nepal | 2011 | DHS | Spain | 2010 | ESS |
| Belgium | 2002 | ESS | Ecuador | 2001 | IPUMS | Ireland | 2002 | ESS | Netherlands | 2002 | ESS | Sweden | 2002 | ESS |
| Belgium | 2010 | ESS | Ecuador | 2010 | IPUMS | Ireland | 2010 | ESS | Netherlands | 2010 | ESS | Sweden | 2010 | ESS |
| Benin | 2001 | DHS | Egypt | 2000 | DHS | Israel | 2002 | ESS | New Zealand | 2000 | ISSP | Switzerland | 2002 | ESS |
| Benin | 2011 | DHS | Egypt | 2008 | DHS | Israel | 2010 | ESS | New Zealand | 2010 | ISSP | Switzerland | 2010 | ESS |
| Bolivia | 1998 | DHS | Estonia | 2004 | ESS | Italy | 2002 | ESS | Nicaragua | 1995 | IPUMS | Tanzania | 1999 | DHS |
| Bolivia | 2008 | DHS | Estonia | 2008 | ESS | Italy | 2012 | ESS | Nicaragua | 2005 | IPUMS | Tanzania | 2009 | DHS |
| Bulgaria | 2000 | ISSP | Ethiopia | 2000 | DHS | Japan | 2000 | ISSP | Niger | 1998 | DHS | Turkey | 2004 | ESS |
| Bulgaria | 2010 | ISSP | Ethiopia | 2005 | DHS | Japan | 2010 | ISSP | Niger | 2012 | DHS | Turkey | 2008 | ESS |
| Burkina Faso | 1998 | DHS | Fiji | 1996 | IPUMS | Jordan | 2002 | DHS | Nigeria | 1999 | DHS | Uganda | 2000 | DHS |
| Burkina Faso | 2010 | DHS | Fiji | 2007 | IPUMS | Jordan | 2009 | DHS | Nigeria | 2010 | DHS | Uganda | 2011 | DHS |
| Cambodia | 1998 | IPUMS | Finland | 2002 | ESS | Kenya | 1998 | DHS | Norway | 2002 | ESS | Ukraine | 2004 | ESS |
| Cambodia | 2008 | IPUMS | Finland | 2010 | ESS | Kenya | 2008 | DHS | Norway | 2010 | ESS | Ukraine | 2010 | ESS |
| Cameroon | 1998 | DHS | France | 2002 | ESS | Korea South | 2003 | ISSP | Panama | 2000 | IPUMS | United Kingdom | 2002 | ESS |
| Cameroon | 2011 | DHS | France | 2010 | ESS | Korea South | 2010 | ISSP | Panama | 2010 | IPUMS | United Kingdom | 2010 | ESS |
| Canada | 2000 | ISSP | Gabon | 2000 | DHS | Kyrgyz Republic | 1997 | DHS | Peru | 2000 | DHS | USA | 2000 | ISSP |
| Canada | 2010 | ISSP | Gabon | 2012 | DHS | Kyrgyz Republic | 2012 | DHS | Peru | 2010 | DHS | USA | 2010 | ISSP |
| Chile | 2000 | ISSP | Germany | 2002 | ESS | Latvia | 2000 | ISSP | Philippines | 1998 | DHS | Uruguay | 1996 | IPUMS |
| Chile | 2010 | ISSP | Germany | 2010 | ESS | Latvia | 2010 | ISSP | Philippines | 2008 | DHS | Uruguay | 2006 | IPUMS |
| Colombia | 2000 | DHS | Ghana | 2000 | IPUMS | Lesotho | 2004 | DHS | Poland | 2002 | ESS | Vietnam | 1999 | IPUMS |
| Colombia | 2010 | DHS | Ghana | 2010 | IPUMS | Lesotho | 2009 | DHS | Poland | 2010 | ESS | Vietnam | 2009 | IPUMS |
| Comoros | 1996 | DHS | Greece | 2002 | ESS | Madagascar | 1997 | DHS | Portugal | 2002 | ESS | Zambia | 2000 | IPUMS |
| Comoros | 2012 | DHS | Greece | 2010 | ESS | Madagascar | 2008 | DHS | Portugal | 2010 | ESS | Zambia | 2010 | IPUMS |
| Costa Rica | 2000 | IPUMS | Guinea | 1999 | DHS | Malawi | 1998 | IPUMS | Russian Federation | 2000 | ISSP | Zimbabwe | 1999 | DHS |
| Costa Rica | 2011 | IPUMS | Guinea | 2012 | DHS | Malawi | 2008 | IPUMS | Russian Federation | 2010 | ISSP | Zimbabwe | 2010 | DHS |

*Appendix 3. Countries included in the regional breakdown of the Barro and Lee dataset*

**Advanced Economies**: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Japan, Luxembourg, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, Turkey, USA, United Kingdom.

**East Asia and the Pacific**: Brunei Darussalam, Cambodia, China, China Hong Kong, China Macao, Fiji, Indonesia, Lao People's Democratic Republic, Malaysia, Mongolia, Myanmar, Papua New Guinea, Philippines, Republic of Korea, Singapore, Taiwan, Thailand, Tonga, Viet Nam.

**Europe and Central Asia**: Albania, Armenia, Bulgaria, Croatia, Czech Republic, Estonia, Hungary, Kazakhstan, Kyrgyzstan, Latvia, Lithuania, Poland, Republic of Moldova, Romania, Russian Federation, Serbia, Slovakia, Slovenia, Tajikistan, Ukraine.

**Latin America and the Caribbean**: Argentina, Barbados, Belize, Bolivia, Brazil, Chile, Colombia, Costa Rica, Cuba, Dominican Rep., Ecuador, El Salvador, Guatemala, Guyana, Haiti, Honduras, Jamaica, Mexico, Nicaragua, Panama, Paraguay, Peru, Trinidad and Tobago, Uruguay, Venezuela.

**Middle East and North Africa**: Algeria, Bahrain, Cyprus, Egypt, Iran (Islamic Republic of), Iraq, Israel, Jordan, Kuwait, Libyan Arab Jamahiriya, Malta, Morocco, Qatar, Saudi Arabia, Syrian Arab Republic, Tunisia, United Arab Emirates, Yemen.

**South Asia**: Afghanistan, Bangladesh, India, Maldives, Nepal, Pakistan, Sri Lanka.

**Sub-Saharan Africa**: Benin, Botswana, Burundi, Cameroon, Central African Republic, Congo Cote d Ivoire, Democratic Republic of the Congo, Gabon, Gambia, Ghana, Kenya, Lesotho, Liberia, Malawi, Mali, Mauritania, Mauritius, Mozambique, Namibia, Niger, Reunion, Rwanda, Senegal, Sierra Leone, South Africa, Sudan, Swaziland, Togo, Uganda, United Republic of Tanzania, Zambia, Zimbabwe.

*Appendix 4. Figures based on relative and absolute Gini coefficients*

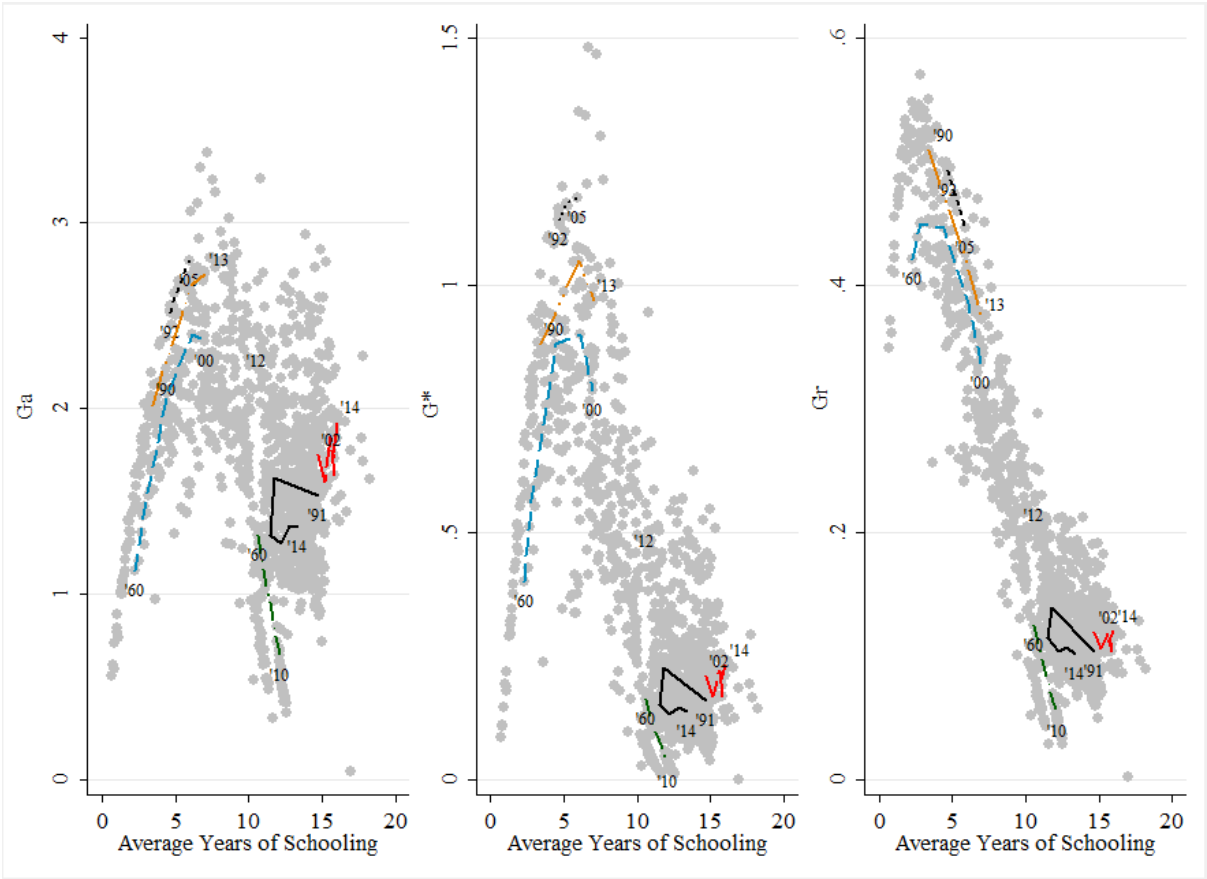


**Figure A1.** Average years of schooling by education inequality within countries using the absolute Gini, G*, and relative Gini coefficients (left, middle, and right panels, respectively). Note: Each dot represents a dataset for a given country and year (1164 datasets); Markers indicate periods; lines based on same data source within each country (e.g. all DHS; all IPUMS; all ISSP; or all ESS). Markers in graph indicate years. Source: Authors' calculations based on IPUMS, ESS, DHS, and ISSP data for individuals aged 30-34.


**Figure A2.** Between-country education inequality in countries' "overall and stage-specific mean years of schooling" according to data source, inequality measure and year. (Cat.): Based on Barro and Lee data for individuals aged 30-34. Ordinal data transformed into years of education. (Cont.): Based on IPUMS; ESS; DHS; and ISSP data for individuals aged 30-34. Countries weighted by population size. Source: Authors' calculations.

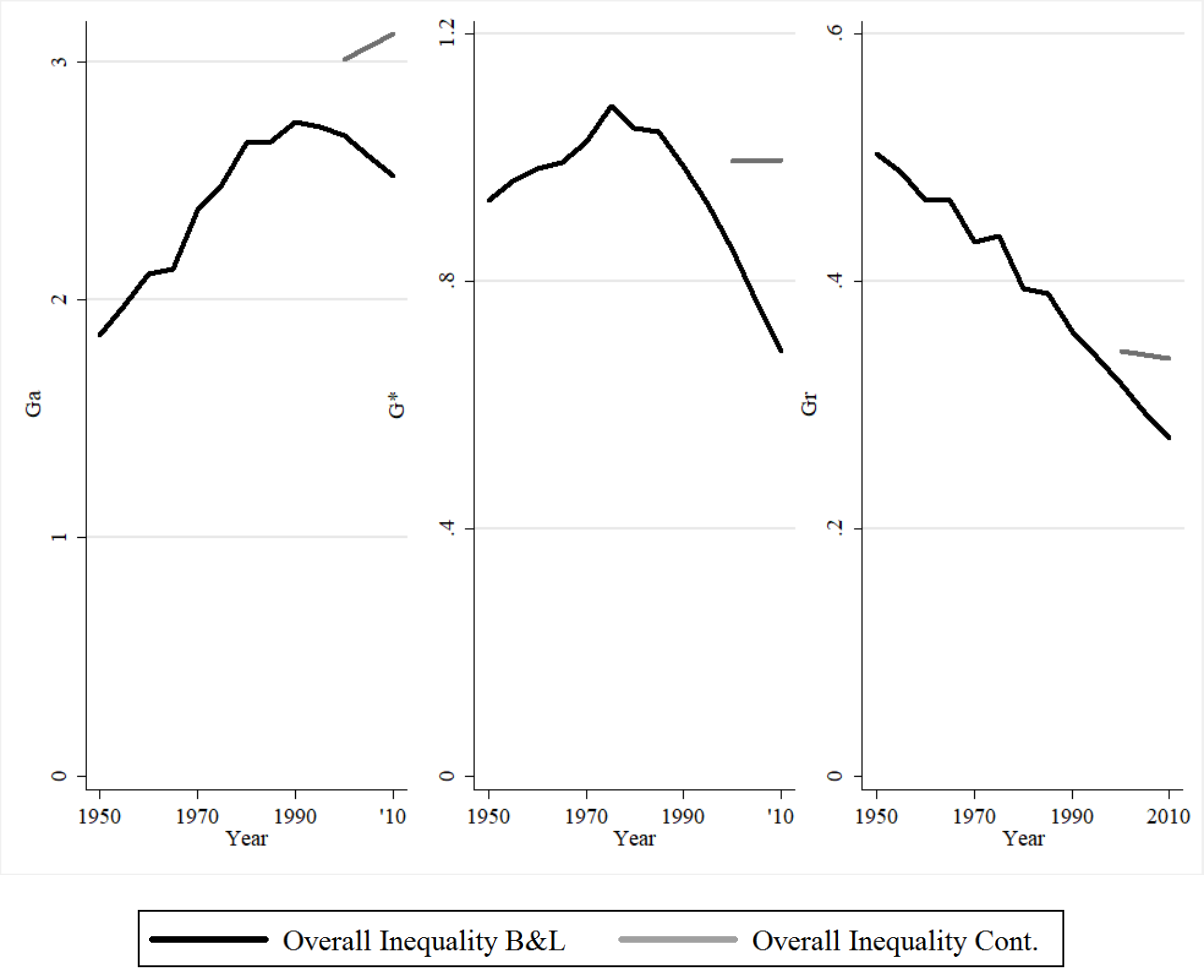


**Figure A3.** Global education inequality according to data source, inequality measure and year. (Cat.): Based on Barro and Lee data for individuals aged 30-34. Ordinal data transformed into years of education. (Cont.): Based on IPUMS; ESS; DHS; and ISSP data for individuals aged 30-34. Source: Authors' calculations.

**Tables and Figures of main text**

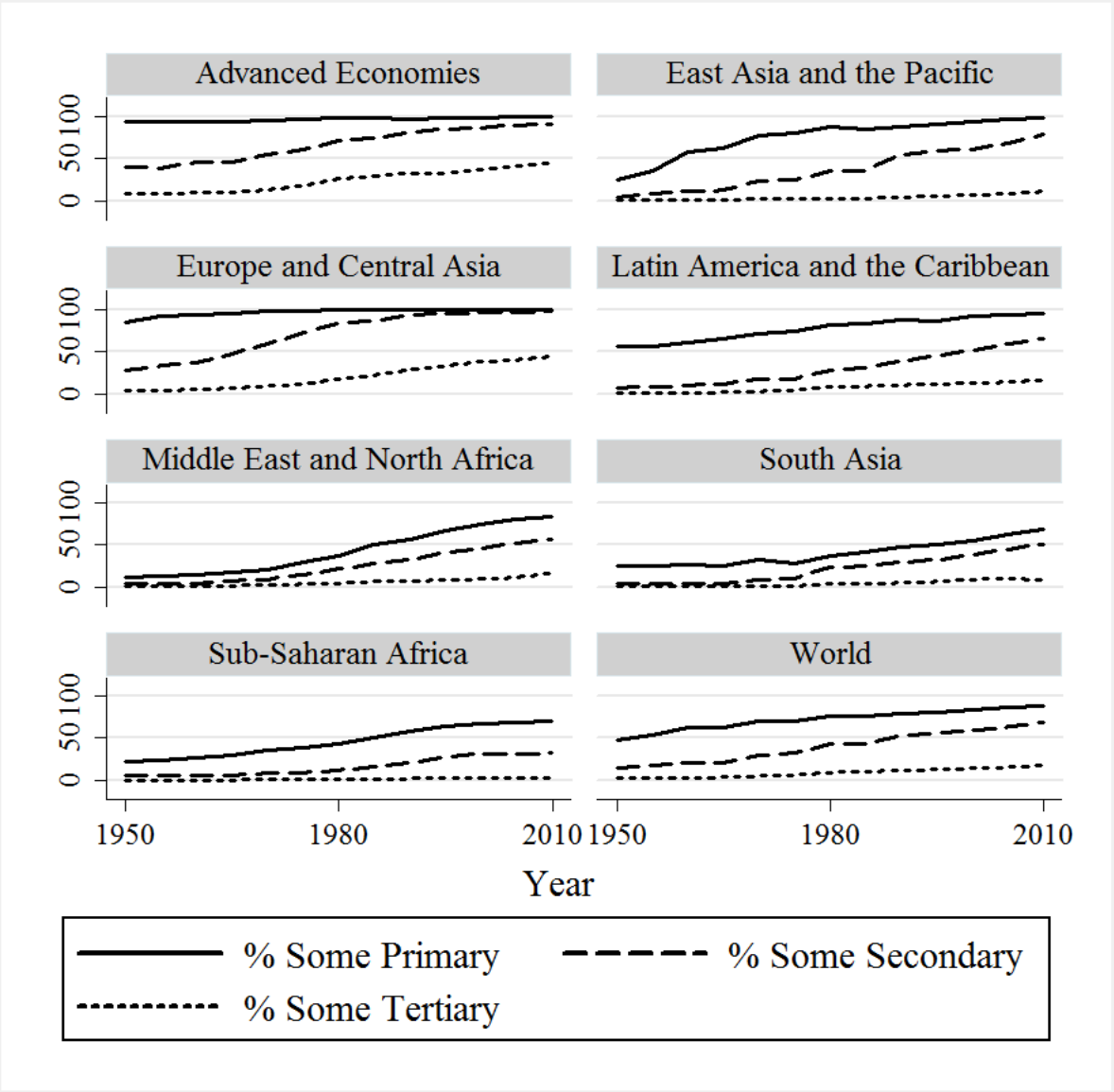


**Figure 1.** Attendance of educational stages 1950-2010 by region for population aged 30-34, weighted by population size of countries. Source: Authors' calculations based on the Barro and Lee dataset.

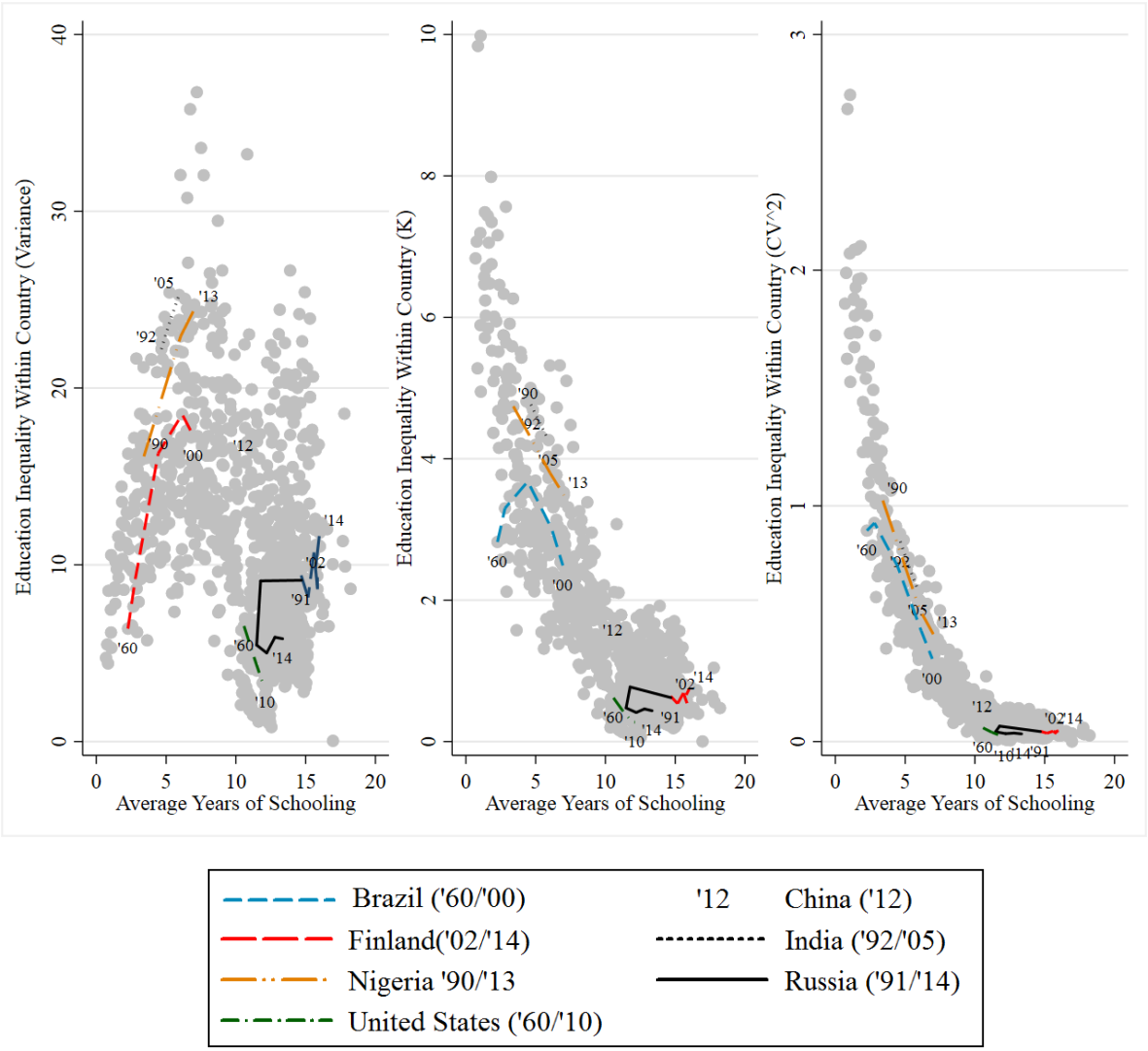


**Figure 2.** Average years of schooling by education inequality within countries using the Variance (left panel), K (middle panel) and the Theil index (right panel). Note: Each dot represents a dataset for a given country and year (1164 datasets); Markers indicate periods; lines based on same data source within each country (e.g. all DHS; all IPUMS; all ISSP; or all ESS). Markers in graph indicate years. Source: Authors' calculations based on IPUMS, ESS, DHS, and ISSP data for individuals aged 30-34.

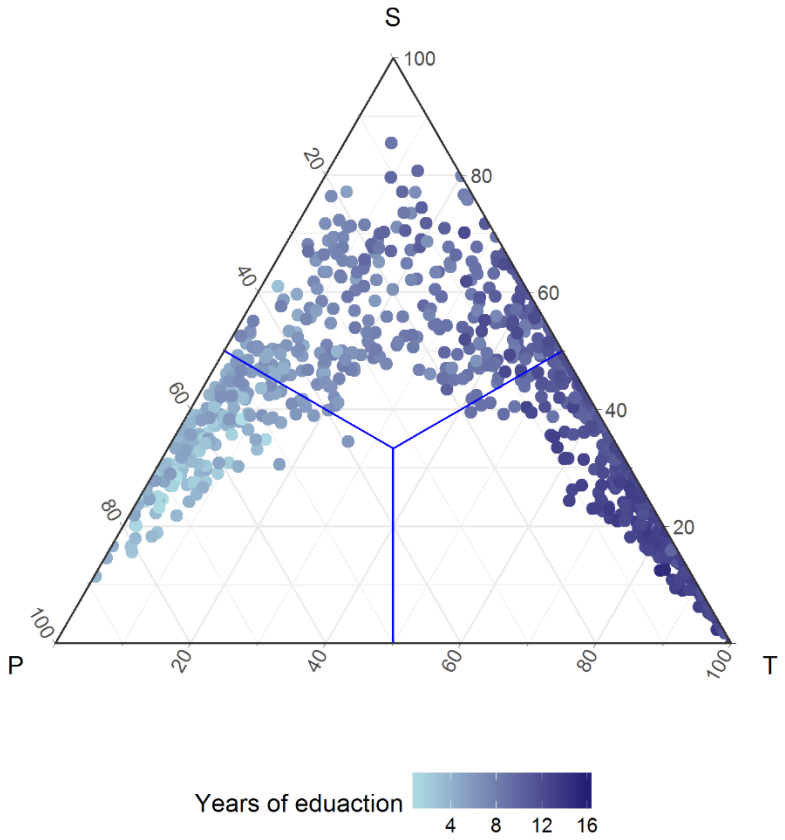


**Figure 3.** Relative contribution of educational stages to overall education inequality within given country-years and by average years of education. Note: P = % of Education inequality due to primary education; S = Secondary; T = Tertiary Education. Colors indicate average years of schooling. Source: Authors' calculations based on IPUMS, ESS, DHS, and ISSP data for individuals aged 30-34.

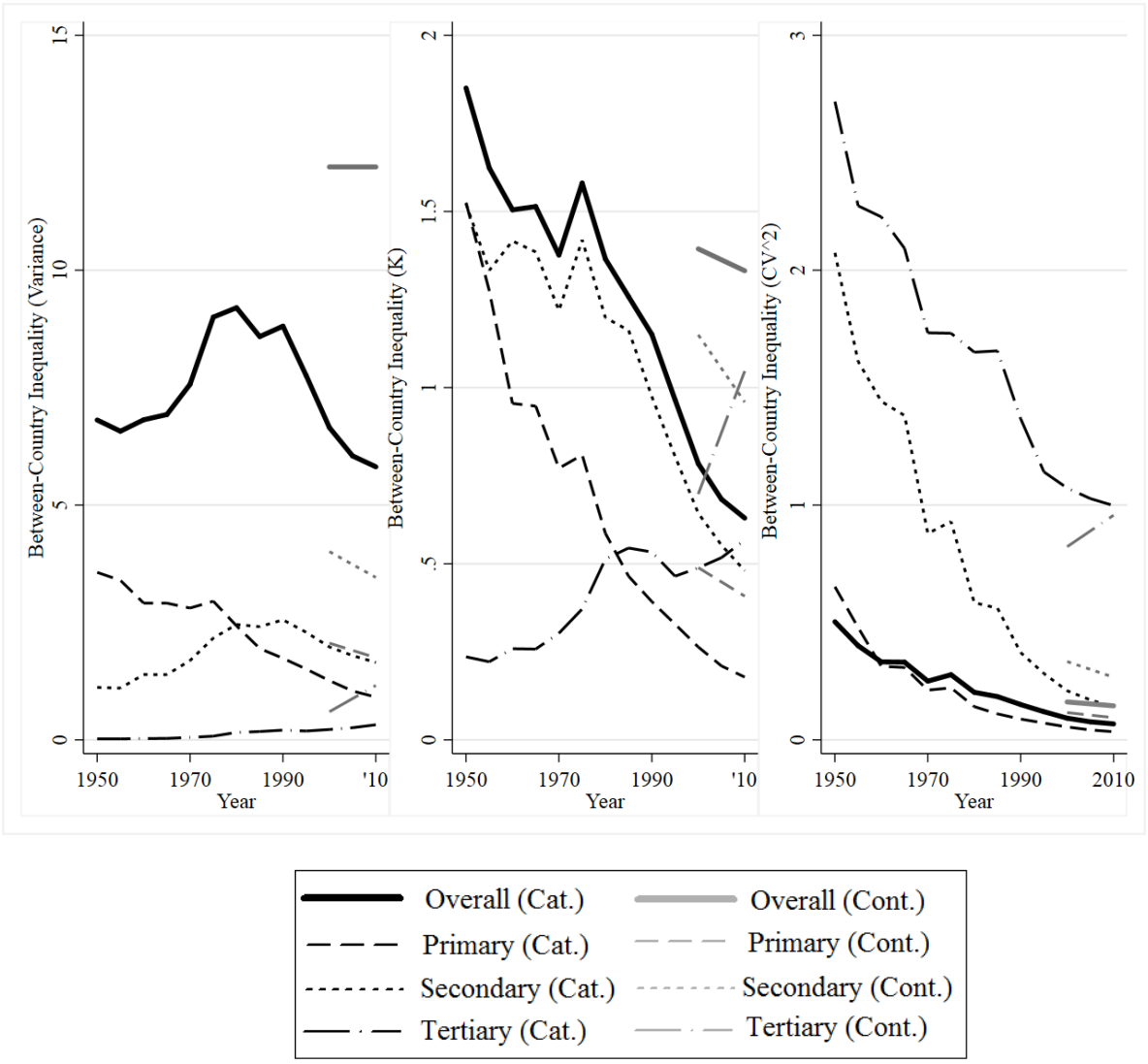


**Figure 4**. Between-country education inequality in countries' "overall and stage-specific mean years of schooling" according to data source, inequality measure and year. (Cat.): Based on Barro and Lee data for individuals aged 30-34. Ordinal data transformed into years of education. (Cont.): Based on IPUMS; ESS; DHS; and ISSP data for individuals aged 30-34. Countries weighted by population size. Source: Authors' calculations.

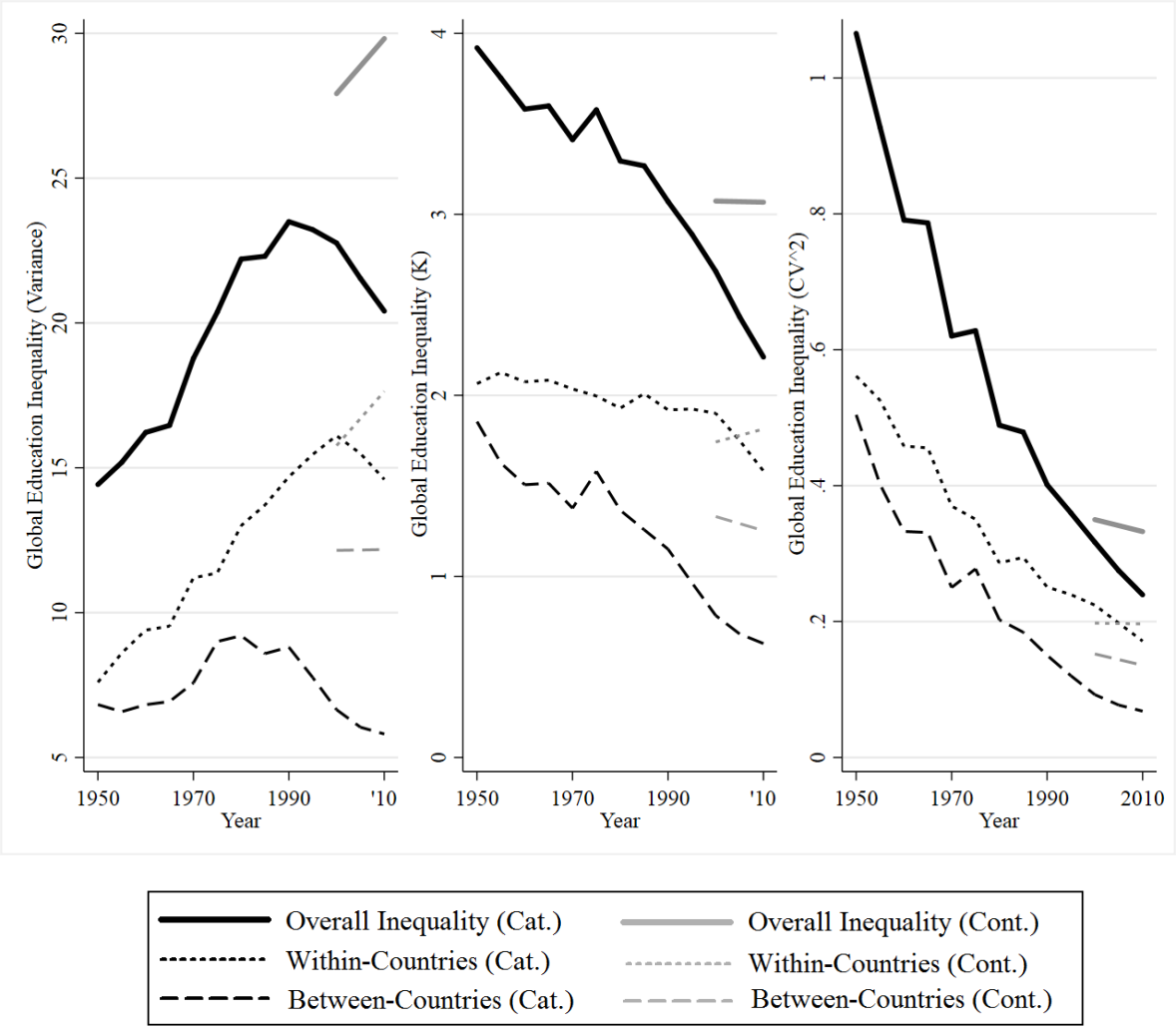


**Figure 5.** Within country, between country and global education inequality according to data source, inequality measure and year. (Cat.): Based on Barro and Lee data for individuals aged 30-34. Ordinal data transformed into years of education. (Cont.): Based on IPUMS; ESS; DHS; and ISSP data for individuals aged 30-34. Source: Authors' calculations. Countries weighted by population size.

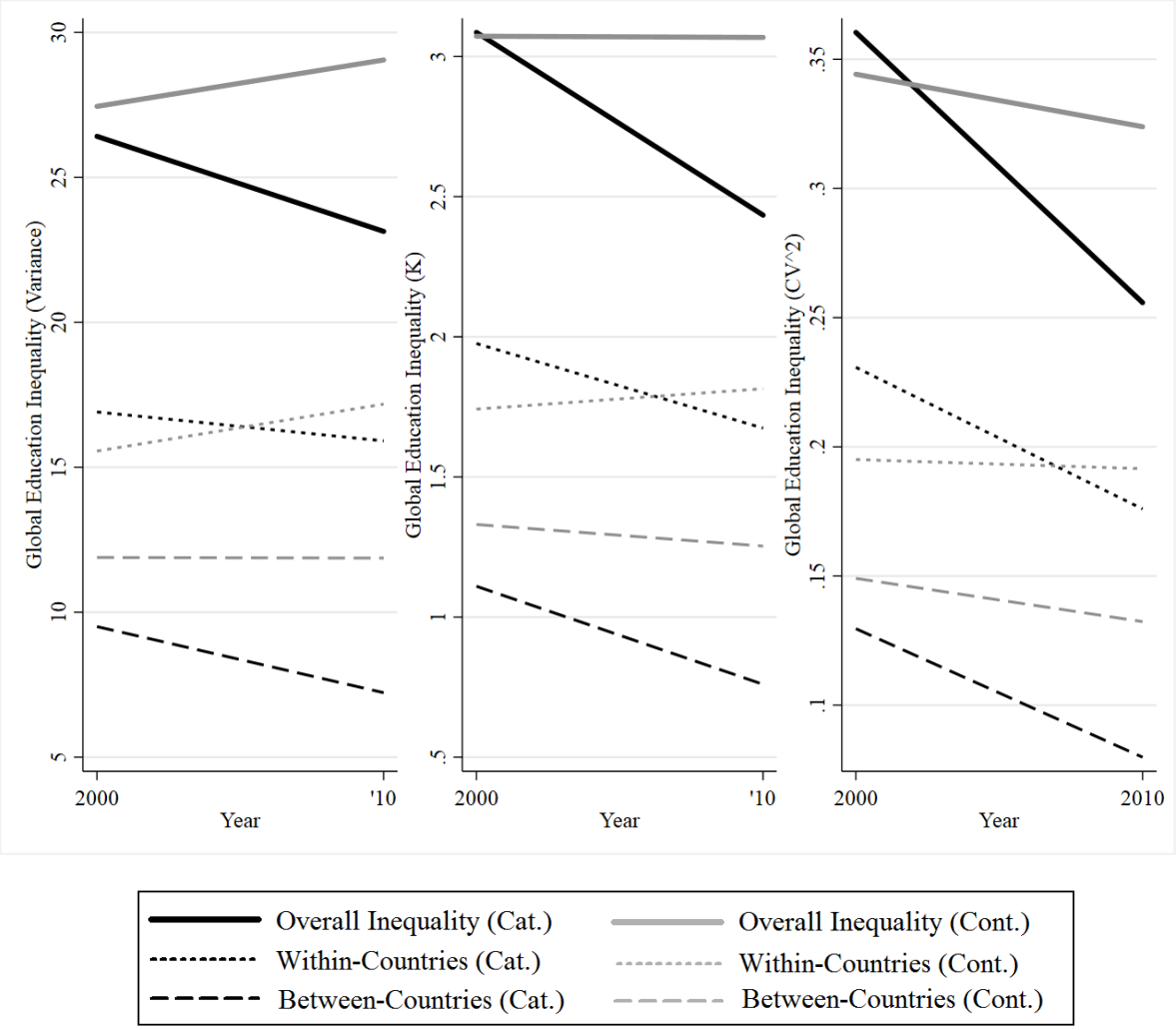


**Figure 6.** Within country, between country and global education inequality according to data source, inequality measure and year, based on the same set of countries present in both datasets. (Cat.): Based on Barro and Lee data for individuals aged 30-34. Ordinal data transformed into years of education. (Cont.): Based on IPUMS; ESS; DHS; and ISSP data for individuals aged 30-34. Countries weighted by population size. Source: Authors' calculations.

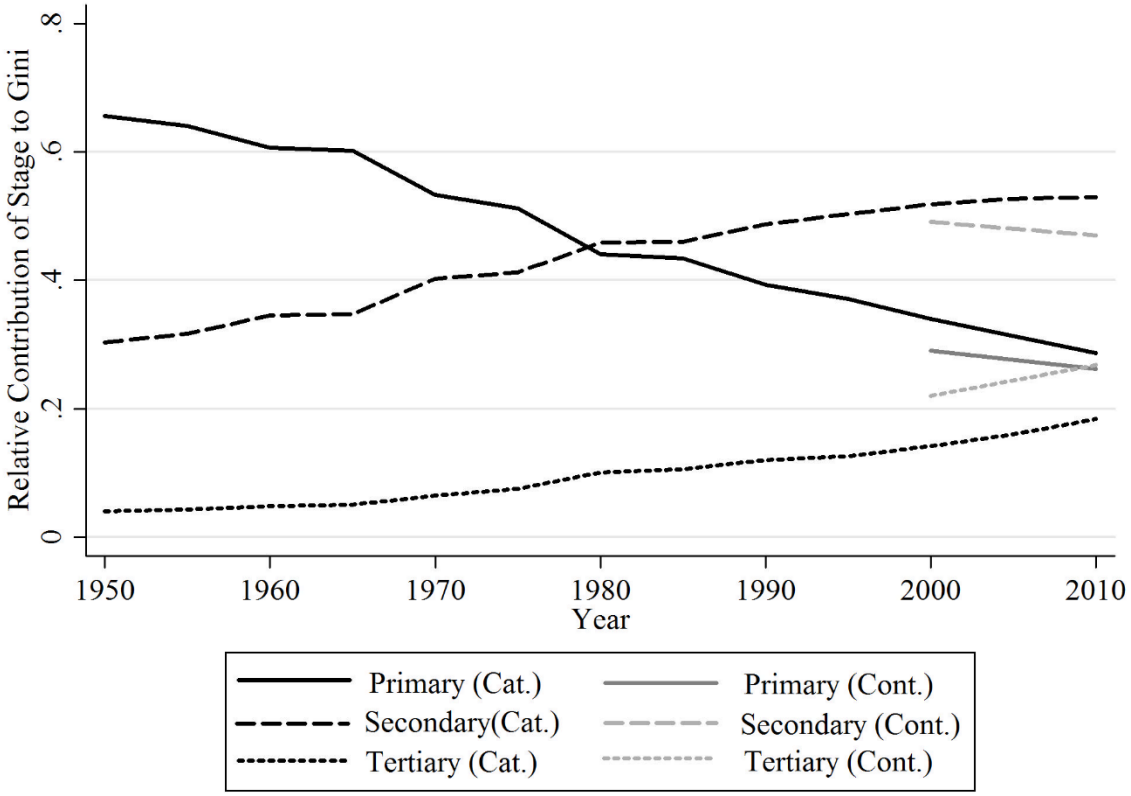


**Figure 7.** Relative contribution of educational stages to global education inequality, depending on data source. (Cat.): Based on Barro and Lee data for individuals aged 30-34. Ordinal data transformed into years of education. (Cont.): Based on IPUMS; ESS; DHS; and ISSP data for individuals aged 30-34. Countries weighted by population size. Source: Authors' calculations.