the review of income and wealth

Review of Income and Wealth
Series 65, Number S1, November 2019
DOI: 10.1111/roiw.12437

THE ROLE OF AGE AND GENDER IN EDUCATIONAL EXPANSION: THE SOUTH ASIAN EXPERIENCE IN THE GLOBAL CONTEXT

BY PETRA SAUER*

Vienna University of Economics and Business

I calculate education Gini coefficients and decompose the overall degree of educational inequality into age, sex, and within-group components for 171 countries from 1970 to 2010. Doing so enables me to analyze the distributional outcomes of educational expansion. I use South Asia as a case study, as the education distribution in the region is among the most unequal in the world. Generally, educational inequality is decreasing over the observed sample period around the globe. Yet, as improvements are initiated by enhancing the educational opportunities of the young, the gap between cohorts widens in transition phases but vanishes thereafter. Gaps between the sexes are reduced substantially, but widen if either males or females are the first to enter higher education levels. Also, gaps within population subgroups follow a similar trajectory. Instead of a Kuznets-curve relation, I thus find evidence for educational inequality to evolve in waves as education expands.

JEL Codes: I24, I25

Keywords: educational attainment, educational inequality, age, gender

1. Introduction

Over the past decades, education has been rapidly expanding around the globe. Not least as a consequence of the United Nations' Millennium Development Goal (MDG) to achieve universal primary education by 2015, educational expansion entails rising population shares with primary and secondary education in low- and middle-income countries. According to Dorius (2012), "we are now fast approaching the time when primary schooling will be a universal fact of life virtually everywhere." In India, for example, 60 percent of the population aged 20–39 did not have any formal education in 1970 and secondary attainment shares were negligible. By 2010, the share of unschooled people had fallen to 26 percent, while primary and secondary attainment shares had increased to 15 percent and 26 percent, respectively.[1] In contrast, the minimum level of education that people in

**Note*: I am grateful to Apoorva Gupta, Bilal Foudad Barakat, two anonymous referees, and colleagues at the Research Institute "Economics of Inequality," as well as participants at the 2018 Canazei Winter School of Inequality and Social Welfare Theory, and at the 2017 IARIW–ICRIER conference on "Experiences and Challenges in Measuring Income, Inequality, and Poverty in South Asia," for helpful comments and discussions.

Correspondence to: Petra Sauer, Centre for Pluralist Economics, Anglia Ruskin University, East Road, Cambridge CB11PT (petra.sauer@anglia.ac.uk) and Research Institute Economics of Inequality (INEQ), Vienna University of Economics and Business (petra.sauer@wu.ac.at).

[1]These figures are obtained from the Wittgenstein Centre Data Explorer.

high-income countries are going for is the upper-secondary level (Ballarino *et al.*, 2013), and tertiary attainment figures have been continuously increasing since the second half of the 20th century (Schofer and Meyer, 2005).[2] Pronounced dynamics in the global education structure raise questions about the distributional consequences of educational expansion. In this regard, the distribution of education is of particular interest. It not only shapes the equalizing impact of rising educational attainment on the income distribution (see, among others, Ballarino *et al.*, 2013; Checchi and van de Werfhorst, 2017; Cruces *et al.*, 2011), but also the extent to which human capital accumulation positively affects the growth prospects of nations (Castelló and Doménech, 2002; Sauer and Zagler, 2012, 2014).

Different degrees of educational inequality across countries over time result from the extent to which policies are able to enlarge the group of people who participate in education. This is done by improving the educational opportunities of women as well as of people from disadvantaged backgrounds. In both respects, providing for enhanced schooling prospects of the young secures improved educational outcomes of future generations. Consequently, different educational expansion trajectories result from different magnitudes of human capital accumulation and equalization among the young, between men and women, and between individuals within demographic groups.

Existing research shows that the distribution of educational attainment within countries becomes more equal as education expands (Castelló and Doménech, 2002; Sauer and Zagler, 2012, 2014). Morrisson and Murtin (2013) and Castelló and Doménech (2002) demonstrate that the strong negative relation between educational inequality and average educational attainment that has been revealed in cross-country comparisons is mechanical and due to the decline in illiteracy. The findings of Sauer and Zagler (2012, 2014) and Meschi and Scervini (2013) provide evidence that a behavioral relationship exists within countries over time. According to Meschi and Scervini (2013), educational inequality has substantially declined in the transition toward universal basic education, but expansion of post-secondary education tends to increase the degree of inequality in the distribution of educational attainment. This contrasts with the hypothesis of an educational Kuznets curve, which implies that inequality should rise before it declines in the process of educational expansion.

By now, studies concerned with the distribution of education have treated all individuals within countries equally. An exception is the analysis of Crespo-Cuaresma *et al.* (2013), who investigate age-group and gender-specific distributions of education. Their findings indicate that educational improvements typically entail rising education levels of the young compared to the elderly and might affect males and females differently. In this paper, I contribute to the literature by providing an integrated analysis of the evolution of gaps within and between demographic subgroups of the population. For a global panel of countries, I thus decompose overall educational inequality into age, gender, and within-group components. Using

[2]However, Ballarino *et al.* (2013) show that this process does not apply equally across high-income countries. While the share of the population with tertiary education approaches 50 percent in the United Kingdom and in Nordic countries, the expansion process slows down at 30 percent in Continental, Mediterranean, and Eastern European countries.

a matrix-algebra approach, Silber (1989) shows how to decompose the Gini index into three components, a within- and a between-group component, and a residual term. I adapt his method so that it can be applied to aggregate education data instead of individual income data. Doing so enables me to separate inequalities that are due to inequality across age groups as well as between males and females from disparities within these groups for 171 countries over the time span from 1970 to 2010, in 5-year intervals. Moreover, I provide an intuitive interpretation of the residual term, which relates to the degree of within-group inequality. I descriptively analyze the evolution of the obtained indicators of educational inequality, putting the experience of South Asian countries into the global context. The use of South Asia for a case study provides interesting insights into the dynamics of educational inequality. Even if strong progress has been made to achieve universal primary education, countries are still characterized by strong divides between women and men, rural and urban areas, and religions, among other things (Dréze and Sen 2013). Regarding the distribution of education between and within population groups, they are thus among the most unequal in the world. Finally, in a panel-regression framework, I test for non-linearities in the relation between average educational attainment, the education Gini, and each of its components.

The remainder of this paper is structured as follows. Section 2 surveys existing studies on the measurement of educational inequality. Thereafter, I use matrix algebra to derive the education Gini coefficient as a measure of between-category inequality and I describe its decomposition in Section 3. I present the data that I use in Section 4. Section 5 provides a snapshot on the particular experience of educational expansion in South Asian countries. Thereafter, I discuss the results of my decomposition analysis in Section 6. In Section 7, I test for non-linearity in the relationship between average educational attainment and the obtained inequality components. I test for robustness to using the Theil index as an additively decomposable inequality measure in Section 8 Finally, Section 9 summarizes and provides pointers for policy as well as further research.

2. Measuring Educational Inequality

The educational attainment of individuals depends, among other things, on the quantity and quality of their formal education, post-school learning, and experience, as well as the informal knowledge existing in their social environment. Quality differentials in developing countries result, for example, from higher pupil–teacher ratios in urban as compared to remote rural areas, which are expensive to serve and tend to be inhabited by minorities (Bing, 2009). Carnoy (2011) discusses discrepancies between private universities, which absorb the bulk of the increasing student numbers, and a small segment of well-funded public universities in Latin America and East Asia. Post-school learning might take place at work or in specialized educational institutions (see, e.g., Schuetze, 2006; Cohn and Addison, 1998), with differing implications for quality and access. In any case, the particular features of education systems and their distributional outcomes differ across countries (Braga *et al.*, 2013). However, it is not possible to observe and measure all aspects of people's educational achievement. Even with data from individual or household surveys, one is almost always restricted to information on formal schooling careers. That is, one observes whether a person did not experience any education or has attained some basic or higher schooling, and one can estimate the number of years associated with the respective education level. From this, it follows that formal schooling is a categorical rather than a continuous variable. It has a lower boundary at zero, an upper boundary given by the duration to complete tertiary education, and categories that correspond to formal education levels.

Two measures have been used primarily in order to investigate the distributional dimension of education.[3] The standard deviation of schooling has been used to explore the impact of the distribution of education on income growth and poverty reduction (e.g. Birdsall and Londono, 1997; López *et al.*, 1998) as well as income inequality (e.g. Lam and Levison, 1991; Inter-American Development Bank, 1999). Furthermore, similar to the concept of income inequality, standard deviations have been applied to test for the existence of an educational Kuznets curve; that is, an inverted U-shape relation between the distribution and the average level of schooling. By relating the standard deviation of 140 countries in 2000 to the average number of years of schooling, Fan *et al.* (2002) confirmed the findings of Londono (1990) and Ram (1990) that educational inequality first increases as the average level of schooling rises and then, after reaching a peak, starts to decline.

However, the standard deviation is only a measure of absolute dispersion and is sensitive to changes in the mean. As a measure of relative inequality, the education Gini coefficient is therefore seen as a more consistent and robust measure of the distribution of education. Some earlier studies (e.g. Rosthal, 1978; Maas and Criel, 1982; Sheret, 1988) used schooling enrollment figures or education finance data for calculating education Gini coefficients for small samples of mostly developing countries. These databases do not, however, accurately reflect the existent stock of human capital. Enrollment ratios are flow variables that add to the future stock of human capital. Even if they constitute an indicator of access to education or equality of opportunity, they do not capture the degree of inequality in educational outcomes. Due to the availability of datasets that, by reporting attainment figures for various education levels, provide a more appropriate picture of the actual distribution of education, more recent studies calculate the education Gini based on the educational attainment of the population concerned. Educational attainment is typically measured by the number of years of schooling achieved, the percentage of individuals who have completed primary, secondary, or tertiary education levels, or by people's actual competencies (Meschi and Scervini, 2014).

As in its application to income inequality, the education Gini coefficient is a measure of mean standardized deviations between all possible pairs of persons and lies in a range between zero and one. A value of zero indicates a perfectly equally distributed education structure, with the opposite being true for a value of one. The former case corresponds to a situation in which the whole population attains the same education

[3]Fan *et al.* (2002) and Meschi and Scervini (2013) also calculate Theil indices of educational attainment and Castelló and Doménech (2002) additionally report the distribution of education by quintiles.

level, irrespective of which level this is. In the latter case, one person completes tertiary education, for example, while the rest of the population does not attain any formal schooling. López *et al*. (1998) were the first to derive Gini coefficients for 12 countries from attainment data. Fan *et al*. (2001) provide a detailed description of the underlying methodology, calculate education Gini's for 85 industrialized and developing countries for the period from 1960 to 1990, and relate them to average educational attainment, educational gender gaps, and real GDP per capita. In their subsequent work (see Fan *et al*., 2002), they further extend the sample to 140 countries from 1960 to 2000. Subsequently, their approach has been utilized for deriving a consistent indicator of the distribution of education, which can be related to income distribution (e.g. Checchi, 2000) and income growth (e.g. Castelló and Doménech, 2002; Sauer and Zagler, 2012, 2014). In contrast to earlier results, plotting Gini coefficients against average educational attainment does not support an education Kuznets curve, but reveals a strong negative relation between the degree of inequality and the average level of educational attainment. However, according to Sauer and Zagler (2012, 2014) and Meschi and Scervini (2013), the relation can be non-linear if changes of educational inequality within countries over time are taken into account.

Crespo-Cuaresma *et al*. (2013) have integrated the demographic dimension into the analysis of educational inequality. They have constructed a dataset of education Gini measures by age group and gender for 175 countries from 1960 to 2010 in 5-year intervals, based on the first version of the IIASA/VID[4] global dataset of populations by age and sex as well as four levels of education.[5] Investigating differential trends of educational inequality within population subgroups, they show that education is more equally distributed among the young than among the elderly, and among men than among women.

In this paper, I build on and add to the analysis of Crespo-Cuaresma *et al*. (2013) and Meschi and Scervini (2013). First, I use the most recent version of the IIASA/VID education dataset, which provides a breakdown of populations into six education categories, incorporating uncompleted primary as well as lower-secondary schooling (see Section 4). Second, I am able to test for non-linearities in the level–inequality relationship using a global sample of 171 countries with a longitudinal dimension. Most importantly, I decompose the overall education Gini coefficient in order to examine the relative relevance of inequality within and between age groups and gender simultaneously.

3. Matrix Algebra and the Education Gini

In an early work, Silber (1989) presents a matrix approach to the computation of the Gini index of income inequality. In the following, I demonstrate how this method is adapted to the categorial structure of aggregate education data. In particular, I derive the Gini coefficient of educational attainment in matrix notation as a measure of between-category inequality. The use of matrix

[4]International Institute for Applied Systems Analysis/Vienna Institute of Demography.

[5]Benaabdelaali *et al*. (2012) have also computed education Gini coefficients by age group based on the Barro and Lee (2010) education dataset.

algebra also enables me to decompose the Gini index by population subgroups, into three components: a within-group component equal to the weighted sum of within-group inequality, a between-group component equal to the weighted sum of between-group inequality, and a residual term that can be interpreted in relation to the ranking of individuals within subgroups.

3.1. *The Education Gini as a Measure of Between-Category Inequality*

According to Silber (1989), for individual data, the Gini index of inequality can be written in matrix notation as

$$I_G^E = e' Gs, \tag{1}$$

where e' is a row vector with n elements equal to $1/n$, n being the number of observed individuals. If educational attainment is measured by the number of years of schooling, one element, s_i, of the column vector s is the share of individual i's attainment in the total number of years of schooling ($Y_T = \sum_{i=1}^{n} Y_i$) in the population concerned. The n elements of s are sorted in descending order according to individual ranks in the education distribution, so that

$$s_1 \geq s_2 \geq \ldots \geq s_i \geq \ldots \geq s_n. \tag{2}$$

The linear operator, introduced by Silber (1989), is the G matrix; which is an $n \times n$ matrix with upper-diagonal elements g_{ij} equal to −1 when $j > i$, lower-diagonal elements equal to 1 when $i > j$, and diagonal elements equal to 0 when $i = j$.

If the available information is limited to the formal duration it takes to complete an education level, individual variation within these categories vanishes. In order to demonstrate that the education Gini is thus reduced to a measure of between-category inequality, I partition the relevant vectors as well as the G matrix. The quantity of subvectors of e and s is given by the number of categories, c, given by the number of education levels. In turn, the quantity of subvector elements depends on the number of individuals, n_h, for which h is the highest education level attained. The partitioned matrix G consists of c^2 submatrices and thus has the following form:

$$\begin{bmatrix} G(n_1,n_1) & \ldots & G(n_1,n_q) & \ldots & G(n_1,n_c) \\ \vdots & \ddots & & & \vdots \\ G(n_p,n_1) & & \ddots & & \vdots \\ \vdots & & & \ddots & \vdots \\ G(n_c,n_1) & \ldots & \ldots & \ldots & G(n_c,n_c) \end{bmatrix}. \tag{3}$$

The main-diagonal submatrices, of dimension $n_h \times n_h$, capture within-category inequality, with zeros in their main diagonal, −1 in their upper-right, and 1 in the lower-left triangle. The submatrices, $G(n_p, n_q)$, for which $q > p$, consist of identical elements equal to −1. If $p > q$, the elements are equal to 1. Summing over partitioned elements, the education Gini can be written as consisting of a within- and a between-category component:

$$I_G^E = \sum_{h=1}^{c} e'(n_h)G(n_h, n_h)s(n_h) + \sum_{p=1}^{c} \left[\sum_{q \neq p}^{c} e'(n_p)G(n_p, n_q)s(n_q) \right] = I_W^E + I_B^E. \tag{4}$$

If no information about within-category variation is available, the within component is redundant, the overall Gini index reduces to its between-category component—that is, $I_G^E = I_B^E$—and the degree of inequality is generally underestimated. Further inspection of I_B^E enables me to show how the between-category measure of educational inequality can be computed using population shares and category averages of numbers of years of schooling.

Defining the share of category h in the total number of years of schooling, $s_{.h} = \sum_{i=1}^{n_h} Y_{ih}/Y_T$, as well as the mean individual share of the number of years in category h, $\bar{s_h} = s_{.h}/n_h$, the between-category contribution of one pq element for which $q > p$ can be written as follows:

$$e'(n_p)G(n_p,n_q)s(n_q) = -(\frac{n_p}{n})n_q\bar{s_q}. \tag{5}$$

In turn, the contribution of an element with $p > q$ looks like the following:

$$e'(n_q)G(n_q,n_p)s(n_p) = \left(\frac{n_q}{n}\right)n_p\bar{s_p}. \tag{6}$$

Overall between-category inequality can be written as the weighted average of its individual pq contributions, I_{pq}^E, with weights equal to the two pertinent categories' populations and years-of-schooling shares, respectively:[6]

$$\begin{aligned} I_B^E &= \sum_{p=1}^{c}\sum_{q>p}^{c}\frac{n_p+n_q}{n}(n_p\bar{s_p}+n_q\bar{s_q})I_{pq}^E \\ &= \sum_{p=1}^{c}\sum_{q>p}^{c}(\frac{n_q}{n})n_p\bar{s_p}-(\frac{n_p}{n})n_q\bar{s_q}. \end{aligned} \tag{7}$$

Both variants of equation (7) use the mean of schooling years in each category in conjunction with population shares. They can thus be easily calculated based on aggregate data of educational attainment. Rearranging the second part of equation (7) enables me to obtain the familiar version of the education Gini index as a weighted sum of differences in category-specific years of schooling (see Section 2):

$$I_G^E = I_B^E = \frac{1}{\bar{Y}}\sum_{p=1}^{c}\sum_{q>p}^{c}(Y_q - Y_p)p_p p_q, \tag{8}$$

where $\bar{Y} = Y_T/n$ is the mean of the number of years of schooling in the population concerned, $Y_h = (\sum_{i=1}^{n_h} Y_i/n_h)/Y_T$ is the average duration it takes to complete education category h, and $p_h = n_h/n$ is the corresponding population share.

[6]See Appendix A, in the Online Supporting Information.

3.2. *Population Subgroups*

In contrast to education categories, individuals from different subgroups of the population cannot be ordered definitely, resulting in overlapping partitions of the education distribution. Nevertheless, the matrix approach provides an intuitive and straightforward method to decompose the education Gini index into population-subgroup contributions.

Following Silber (1989), I define an additional partitioned vector, v, which is ordered, first, by subgroup averages of education attainment shares and, second, by individual attainment shares within subgroups. Thus,

$$\bar{v}_1 \geq \ldots \geq \bar{v}_j \geq \ldots \geq \bar{v}_g \tag{9}$$

and

$$v_{1,j} \geq v_{2,j} \geq \ldots \geq v_{n_j,j} \quad \forall j, \tag{10}$$

where $v_{i,j}$ is the educational attainment share of individual i in group j and $\bar{v}_j$ is the groups' mean attainment share. The number of subvectors depends on the number of groups g, and the quantity of elements, n_j, varies according to the group's population size. Also, if e and G are partitioned by population subgroup and defined just as in Section 3.1, a modified inequality index, $e'Gv$, can be decomposed into a within-group and a between-group component:

$$\sum_{j=1}^{g} e'(n_j)G(n_j,n_j)v(n_j) + \sum_{a=1}^{g}\left[\sum_{b\neq a}^{g} e'(n_a)G(n_a,n_b)v(n_b)\right] = I_W^E + I_B^E. \tag{11}$$

Rewriting the within-group elements in terms of group attainment and population shares enables me to obtain the within-group component as a weighted average of within-group inequality, with weights equal to the groups' population and attainment shares ($v_{.j}$), respectively. Thus,

$$I_W^E = \sum_{j=1}^{g} \frac{n_j}{n} v_{.j} I_{Gj}^E, \tag{12}$$

where the within-group inequality, I_{Gj}^E, is the between-category Gini index of subgroup j. Beyond that, in a similar manner as in Section 3.1, the between-group component of the education Gini can be shown to be equal to a weighted average of pairwise contributions, I_{ab}^E, with weights equal to the two groups' joint population and attainment shares:

$$\begin{aligned} I_B^E &= \sum_{a=1}^{g}\sum_{b>a}^{g} \frac{n_a+n_b}{n}(n_a\bar{v}_a + n_b\bar{v}_b)I_{ab}^E \\ &= \sum_{a=1}^{g}\sum_{b>a}^{g} \left(\frac{n_b}{n}\right) n_a\bar{v}_a - \left(\frac{n_a}{n}\right) n_b\bar{v}_b. \end{aligned} \tag{13}$$

Defining the mean number of years of schooling in subgroup j as $\bar{Y}_j = \sum_{i=1}^{n_j} Y_{ij}/n_j$, the second part of equation (13) can be rearranged in order to obtain the between-group contribution as a weighted sum of differences in the subgroups' mean number of years of schooling:

$$I_B^E = \frac{1}{\bar{Y}} \sum_{a=1}^{g} \sum_{b>a}^{g} (\bar{Y}_a - \bar{Y}_b) p_a p_b. \quad (14)$$

Finally, the difference between the inequality measures obtained from using the definitely ordered versus the reordered attainment share vectors, $e'Gd = e'G(s-v) = e'Gs - e'Gv$, builds the third component of the Gini index decomposition. This factor can be interpreted as the intensity of modifications necessary to rank individuals according to their groups' educational attainment, or as the degree to which groups are overlapping.

4. Data

In order to compute education Gini coefficients according to equation (8) and decompose them by age and gender, I require information about the educational structure of populations and the typical numbers of years that people spend in school to attain an educational level. I obtain the full distribution of educational attainment over six categories by 5-year age groups and gender from data provided by the Wittgenstein Centre for Demography and Global Human Capital (WIC)[7]. To assemble this dataset, Goujon *et al.* (2016) collect baseline data for 2010 from censuses if available, or from surveys otherwise, and apply harmonization procedures[8] to obtain six education categories that correspond to UNESCO's International Standard Classification of Education (ISCED 1997) as summarized in Table 1. Goujon *et al.* (2016) subsequently adopt the demographic method of multistate projection, using education-specific estimates of fertility, mortality, and migration in order to project from 2010 backward to 1970 and forward to 2060.[9] I use the historical part of their data to compute educational attainment shares for the total population aged 15, each of 18 5-year age groups, and both males and females.

As shown in Table 1, education categories 3–5 are composed of individuals who complete the respective but not the subsequent education level. For example, people in category 3 not only go through primary but also through some lower-secondary education. Potancoková *et al.* (2013) adjust data on formal schooling cycles from the UNESCO Institute for Statistics (UIS) to allow for incomplete levels using region-specific correction factors based on survey estimates. They also provide

[7]I thank Samir K.C., Anne Goujon, and Michaela Potancoková for providing this dataset. A public version is available via http://dataexplorer.wittgensteincentre.org/shiny/wic/ (last accessed November 13, 2018).

[8]Harmonization is necessary in order to allocate country-specific educational categorizations to the ISCED 1997 categories, and to differentiate between completed and incomplete levels. Detailed information about the methodology can be found in Bauer *et al.* (2012).

[9]Multistate population projection methodology uses the characteristic that, once attained, the level of formal education generally does not change over the life course of individuals. Thus, the education structure in any time t can be translated into earlier ($t-n$) or later ($t+n$) periods, applying assumptions on mortality and migration differentials and the education transitions (Goujon *et al.*, 2016).

TABLE 1
Categories of Educational Attainment

	Category[a]	ISCED 1997 level
1	No education	No level of ISCED0 Grade 1 of ISCED 1 not completed
2	Incomplete primary	Incomplete ISCED1
3	Primary	Completed ISCED1 Incomplete ISCED2
4	Lower secondary	Completed ISCED2 Incomplete ISCED3
5	Upper secondary	Completed ISCED3 Incomplete ISCED4 or 5B
6	Post-secondary	ISCED 4 nd 5B[b], ISCED 5A and 6[c]

[a]See Goujon *et al.* (2016). [b]First diploma, shorter post-secondary courses. [c]Longer post-secondary courses, postgraduate level.

estimates of the average number of years of incomplete primary education but apply a constant of 4 years for post-secondary education to balance the diverseness of this category. I use their adjusted country- and cohort-specific duration data to measure the typical duration of each education level in the computation of the average number of years of schooling and education Gini coefficients (see equation 8).

I obtain a dataset with information on average educational attainment and educational inequality for the total population aged 15 and over, as well as for each of 18 5-year age groups, and both males and females. I cover 171 countries over the time period from 1970 to 2010 in 5-year intervals. Figure 1 presents the resulting structure for each country–year observation of the data for the example of India in 2010. The average educational attainment of 5.49 years of schooling and the corresponding education Gini, equal to 0.54, mask differences between demographic subgroups. Plotting both measures against age gives rise to downward- and upward-sloping curves for the average level and inequality, respectively. This indicates that educational improvements start among young people. If these not only affect the average level but also the spread, education is more equally distributed among the young than among the elderly. Even if these relations holds globally, countries differ with respect to the slopes; that is, the magnitude of educational expansion and equalization.[10] Moreover, they differ with regard to the gender gap that exists not only for the average level of educational attainment but also for its distribution. In most countries, education is more equally distributed among men than among women, at least for older cohorts. The polarization between a small segment of highly educated people and a broad group with basic education is thus more pronounced among women than among men. Differences between men and women tend to be lower or vanishing for younger cohorts. However, inequalities below the age of 25 can be substantially underestimated, as some individuals have not completed post-secondary education. In India, the gender gap in education (inequality) is relatively large, with 2.5 years (two Gini points) for cohorts older than 45. Interestingly, the gender gap in average educational attainment diminishes again for cohorts older than 70.

[10]For a detailed discussion and analysis of the dynamics of age-group specific educational inequality, see Crespo-Cuaresma *et al.* (2013).

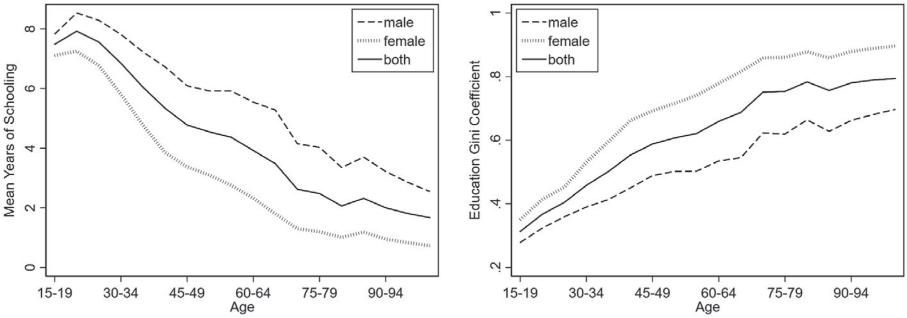

Figure 1. Average Attainment and Educational Inequality, India, 2010 [Colour figure can be viewed at wileyonlinelibrary.com]

In order to ascertain to measure the degree of inequality in completed educational attainment, I restrict my analysis to cohorts aged 25 and over. The finer the grouping, the more homogeneous the groups are. This increases the relative importance of between- versus within-group variation. Considering this impact on the Gini index decomposition, I construct three broader age groups for people aged 25–44, 45–64, and 65 and over. I thus obtain six subgroups of the population, which I use to examine the relevance of age, gender, and within-group differences for the degree of inequality in educational attainment.

5. South Asia in the Global Context

Even if South Asian countries report modest degrees of inequality in monetary terms as compared to countries at a similar stage of development, inequalities in human development outcomes are among the highest in the world (Rama *et al*., 2015). This implies that capabilities to transform resources into individual well-being are unequally distributed among South Asian populations. Yet, the region has followed the global trend of educational expansion, resulting in substantial improvements in access to elementary education. Gaps in educational attainment thus tend to be narrower among children than among adults (Rama et al., 2015). But educational outcomes vary substantially across countries within the region. Looking at the educational attainment of the richest relative to the poorest quintile, Rama *et al*. (2015) show that the region consists of countries with the most equal (Sri Lanka and the Maldives) and the most unequal (Nepal, Afghanistan, and Bhutan) distributions of education in the developing world.[11]

[11]In this analysis, I am not able to capture heterogeneity within countries, which can be equally pronounced. In India, for example, the state of Kerala, with its long tradition of pro-education policy, is much more similar to Sri Lanka than to the northern states of India, which lag behind in their educational achievements.

Consideration of the distribution of educational attainment measured by the education Gini coefficient reveals that countries in South Asia (SA)[12] are among the most unequal in the world. The average education Gini (0.54) of the region is not only higher than the global average in 2010 (0.33), but also larger than the respective value for sub-Saharan Africa (SSA) (see Table 2). Both regions also report the lowest numbers of mean years of schooling. In contrast, high-income (HI) OECD countries, as well as countries in Central Asia and Eastern Europe (CAE), show the lowest Gini values in conjunction with the highest average level of educational attainment, on average. The latter region was able to reduce educational inequality by 63 percent between 1970 and 2010, the change being mainly driven by the significant reduction in educational inequality in the southeastern part of Europe.[13] With a change of 52 percent, the general trend toward more equal education distributions was also particularly pronounced in countries of the Middle East and North Africa (MENA), which more than doubled their average educational attainment. Substantial improvements also took place in East Asia and the Pacific (EAP), which is in the medium spectrum of global educational inequality and even reports a slightly lower level than countries in Latin America and the Caribbean (LAC) in 2010. The decline over time is not only sluggish in HI OECD countries, where the initial level was already low, but also in persistently high-inequality countries in SA and SSA. Yet, South Asian countries experienced the strongest increase in mean numbers of years of schooling, on average, over the observed sample period. This suggests that educational expansion has not enabled educational divides in the region to be substantially overcome, and that gaps between and within demographic groups are important factors in shaping overall inequality.

Access to primary education has been recognized as a basic human right and a prerequisite for equality of opportunity in most countries of South Asia. The goal of universal primary education has thus triggered strong progress in coverage as well as equality (Rama *et al.*, 2015). Across the region, net attendance and completion rates are thus highest at the primary level, and indicate almost full coverage in Bangladesh, the Maldives, and Sri Lanka (see columns (1) and (2) of Table 3). However, the quality of schools has not kept pace with expansion at the primary level (Dréze and Sen, 2013). Existing studies for South Asian countries provide evidence that learning outcomes are low on average, and are unequally distributed across social groups (e.g. World Bank, 2013; Harsha *et al.*, 2013). Major explanatory factors are the lack of qualified teachers and teacher absenteeism (Dréze and

[12]Unfortunately, IIASA/VID data is not available for Afghanistan and Sri Lanka. Thus, neither country is included in this analysis. For a list of countries by world region, see Appendix B. The country grouping follows the geographical regions defined by the World Bank (https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups; last accessed November 16, 2018), with the exception that I separate OECD countries that the World Bank classifies as high-income economies.

[13]The group of CAE countries consists of countries in Eastern Europe (e.g. Latvia and Ukraine), in Central Asia (e.g. Armenia and Kazakhstan), as well as countries in southeastern Europe (e.g. Albania, Macedonia, and Turkey). While educational inequality is historically relatively low in the former two groups of countries, the latter group had relatively high levels of educational inequality in the 1970s and 1980s and was able to reduce it significantly thereafter.

TABLE 2
Summary Statistics 1

Region[a]	$\bar{Y}^b$	$\Delta\bar{Y}^c$	I_G^E	ΔI_G^E
CAE	10.33	79.81	0.16	−62.89
HI OECD	11.74	37.41	0.16	−36.41
EAP	8.24	127.85	0.29	−47.64
LAC	8.36	84.10	0.30	−38.06
MENA	8.00	222.84	0.36	−51.52
SSA	4.79	257.93	0.54	−33.76
SA	4.36	259.16	0.59	−30.17
Global	8.21	91.56	0.33	−41.04

[a]CAE, Central Asia and Eastern Europe; EAP, East Asia and the Pacific; HI OECD, high-income OECD; LAC, Latin America and the Caribbean; MENA, Middle East and North Africa; SA, South Asia; SSA, sub-Saharan Africa. [b]Levels are from 2010. [c]Total change between 1970 and 2010, in percent.

Sen, 2013). According to the findings of Kremer *et al.* (2005), fewer than half of the teachers employed in 3000 randomly selected Indian schools were actively teaching on an average day. Such supply-side factors, in conjunction with issues of security, the availability of latrines and warm meals at schools, among other things, have been shown to inhibit progression of pupils from basic to higher education (Dréze and Sen, 2013) and to disproportionately affect girls' educational attainment (Swainson, 2006). Inequalities with respect to the quantity of education are thus still relevant at higher levels of education. Hence, although differences across countries exist, net attendance and completion rates decline substantially with increasing education levels (see Table 3).[14] At the same time, gaps in completion rates that are due to gender, whether children go to school in rural or urban areas, or whether they grow up in poor or rich households become larger when moving from primary (p) to lower- (ls) and upper-secondary education (see Table 4). At the upper-secondary level, the relative advantage of girls in Bangladesh, Bhutan, and the Maldives is reversed. It is only in Sri Lanka that girls increasingly outperform boys. Table 4 also provides indication that differences between individuals within population groups are due, among other things, to gaps between urban and rural areas, and even more so between children from different economic backgrounds. Similarly, (Rama *et al.*, 2015) provide evidence showing that parental education, gender, location, region, caste, and religion contribute significantly to inequality in primary, and particularly in lower-secondary, attainment.

[14]Net attendance rates are measured as the share of students in the theoretical age group for a given level of education attending that level compared to the total population in that age group. Attendance rates thus capture access and coverage. Completion rates, on the other hand, indicate how many persons in a given age group have completed the pertinent education level without excessive delay (see the UIS data information at data.uis.unesco.org; last accessed November 22, 2018). They are given as the percentage of persons aged 3–5 years above the intended age for the last grade of each level who have completed that grade. Completion rates can be higher than attendance rates if the share of over-age students and repetitors is large. If, for example, 25 percent of the population enters primary education 1 year after the formal entrance year, the attendance rate is depressed to 93.75 percent, even if all of them complete. The results in Table 3 show that this is particularly relevant for the lower-secondary level in all countries. However, while the difference is substantial in India, Sri Lanka achieves a timely transition from primary to secondary education for the majority of its young population.

TABLE 3
Educational Coverage in South Asia

	Primary		Lower secondary		Upper secondary		
Country	Att.[b]	Comp.[c]	Att.	Comp.	Att.	Comp.	Tertiary[a]
Afghanistan	60.03	54.21	31.19	36.96	22.95	24.08	12.58
Bangladesh	85.68	80.01	46.32	54.93	36.44	18.90	16.25
Bhutan	89.61	67.86	45.29	38.78	18.61	20.98	3.26
India	79.36	91.58	55.52	80.84	54.14	42.89	—
Maldives	91.92	96.78	68.68	77.87	10.77	13.21	3.31
Nepal	75.08	76.23	43.35	65.57	41.88	—	21.01
Pakistan	62.08	60.90	32.86	45.57	28.01	19.51	12.52
Sri Lanka	87.76	98.38	83.43	88.11	58.44	25.03	0.03

[a]For tertiary education, only gross attendance rates are available, which are computed by dividing the number of students attending a given level of education regardless of age by the population of the age group that officially corresponds to the given level of education. [b]Net attendance rate. [c]Completion rate. *Data Source*: UIS. Latest figures for each country: Afghanistan, 2015; Bangladesh, 2014; Bhutan, 2010; India, 2016; Maldives, 2008; Nepal, 2016; Pakistan, 2012; Sri Lanka, 2006.

6. Decomposing Educational Inequality

Differences in the distributional effects of educational expansion across countries might stem from the extent of equalization between males and females, and between individuals within demographic subgroups of the population. But educational expansion can enlarge gaps between younger and older cohorts. In order to provide an integrated picture of differences between age groups and sexes, as well as differences within age–gender groups, I apply equations (8), (12) and (14) to decompose the education Gini coefficient of the total population aged 25 and over of 171 countries from 1970 to 2010 into five components as follows:

$$I_G^E = I_{\text{age}}^E + I_{\text{age}/sex}^E + I_{\text{sex}}^E + I_{\text{within}}^E + I_{\text{residual}}^E, \tag{15}$$

where I_{age}^E captures the contribution of gender-specific differentials between age groups, and can be further decomposed into male (I_{ageM}^E) and female (I_{ageF}^E) components. $I_{\text{age}/sex}^E$ compares the educational attainment of different sexes *and* cohorts. I_{sex}^E is the component of the education Gini that is due to differentials between males and females of identical age groups. The extent of inequality within population subgroups, in turn, is given by I_{within}^E. Finally, I_{residual}^E is the residual component.

The upper panel of Figure 2 shows how educational inequality and each component evolved over time in South Asian countries. Differences in levels and changes of the education Gini reveal the strong equalizing impact of expansion of lower education levels in the Maldives and Bangladesh, while Bhutan and Pakistan, which have lower-primary and lower-secondary completion rates, lag behind (see Table 3). Beyond that, the decline in overall inequality across all countries of the region involves declining between- and within-group inequality.[15] However, the relative rel-

[15]See below for a discussion about the residual term.

TABLE 4
Parity Indices

	Female/Male			Rural/Urban			Poorest/Richest[a]		
Country	p[b]	ls	us	p	ls	us	p	ls	us
Afghanistan	0.60	0.52	0.45	0.68	0.56	0.46	0.60	0.41	0.29
Bangladesh	1.11	1.02	0.82	0.99	0.94	0.61	0.70	0.40	0.10
Bhutan	1.06	0.94	0.73	0.67	0.46	0.42	0.47	0.21	0.11
India	1.00	0.96	0.85	0.97	0.92	0.65	0.82	0.62	0.18
Maldives	1.03	1.13	0.73	0.97	0.83	0.30	0.95	0.72	0.17
Nepal	0.98	0.92		0.85	0.76		0.81	0.60	
Pakistan	0.92	0.82	0.85	0.71	0.62	0.42	0.27	0.14	0.07
Sri Lanka	1.01	1.05	1.33	1.00	1.02	0.70	0.97	0.80	0.18

[a]Referring to quintiles of the wealth distribution. [b]p, Primary; ls, lower secondary; us, upper secondary. Adjusted parity indices are based on the ratio, for example, between female and male completion rates, and adjusted to be symmetrical around 1 and to rule out extensively high values. *Data Source*: UIS. Latest figures for each country: Afghanistan, 2015; Bangladesh, 2014; Bhutan, 2010; India, 2016; Maldives, 2008; Nepal, 2016; Pakistan, 2012; Sri Lanka, 2006.

evance of age, gender, and within-group inequality varies over time and across countries. The lower panel of Figure 2 plots each component's share in total inequality over time. In general, the reduction of differences between females and males has been the most significant factor, which contributes to declining educational inequality. Nevertheless, the contribution of gender inequality was relatively large (greater than 20 percent) until the 1990s in Bhutan and Nepal. Since the beginning of the new millennium, gender inequality has improved considerably in both countries. Pakistan started at a lower level in 1970, but differences between males and females reduced only marginally. The slow decline in overall educational inequality in these countries is thus partly explained by the sluggish reduction in gender inequality. In India, educational gaps between males and females made up a constantly declining share in total inequality (from 16 percent in 1970 to 9 percent in 2010). Such a development was particularly pronounced in Bangladesh, where the respective ratio declined from 19 percent in 1970 to 5 percent in 2010. In contrast, the relative relevance of gender inequality has already been low since the 1970s in the Maldives but decreased further by 4 percentage points to 1.5 percent in 2010.

The contribution of within-group inequality is relatively constant around the regional average of 24 percent in each country. I find a significant declining time trend[16] for the whole region, which is mainly driven by improvements in the Maldives and Nepal. Yet, in neither country did the share fall below 20 percent over the observed time period and, after declining, the relative contribution of within-group inequality has increased in Bangladesh since 1985 and in Bhutan since 2005. This implies that educational divides within demographic subgroups of the population continue to exert a significant influence on the total level of educational inequality.

As educational improvements start among young people, gaps between age groups widen in transitional phases. Thus, the relative contributions of both age

[16]This is obtained from a regression of the relevant component (contribution) on time, controlling for country-fixed effects.

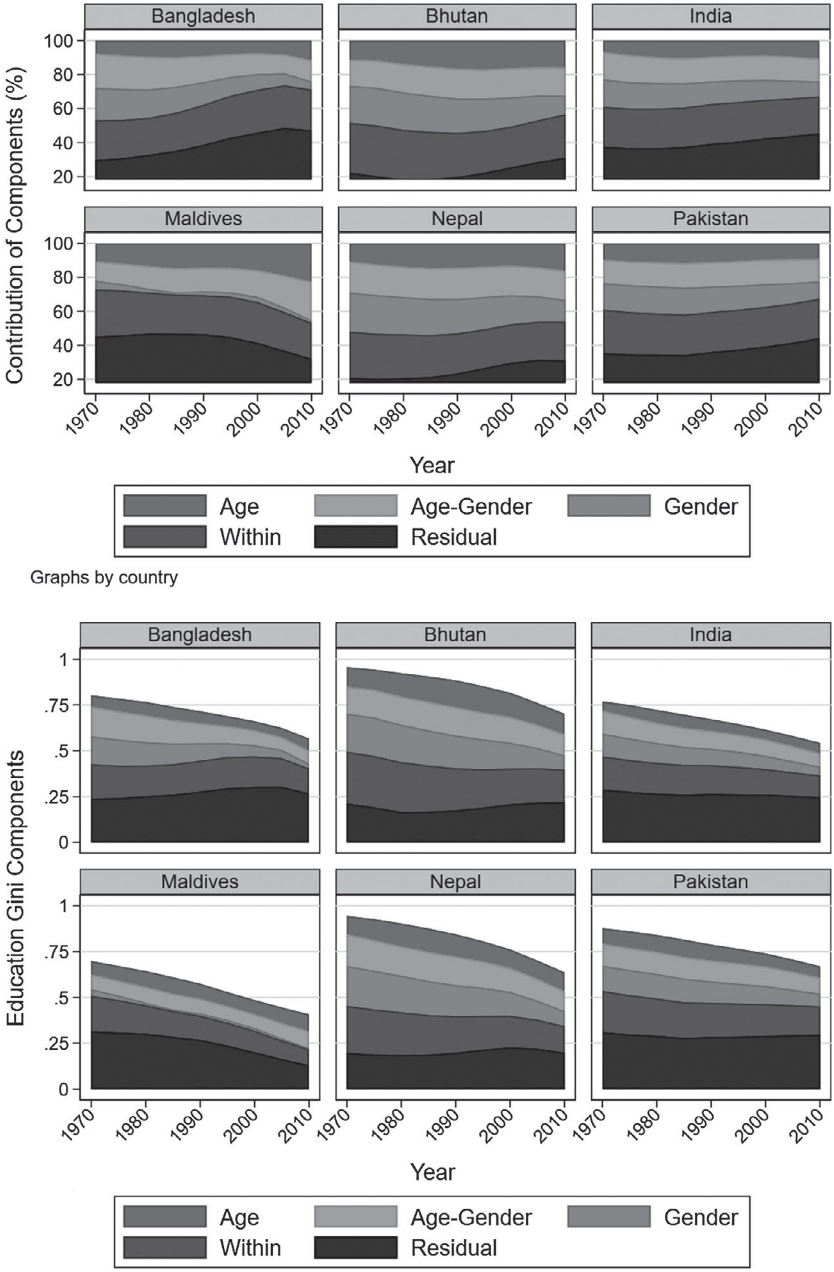

Figure 2. Educational Inequality Components in South Asia

components increase as the relative relevance of gender and within-group inequality declines, particularly in the Maldives, but also more recently in Bangladesh, Bhutan, India, and Nepal.

The developments in South Asian countries already hint at more general findings concerning the relationship of within- and between-group inequality with overall educational inequality and educational expansion. To put the experience of South Asia into the global context, I compare world regions in Table 5 and investigate the relation between the education Gini of the total population aged 25 and over with each of its components, using the global sample of countries in Figure 3.

In 2010, the contribution of the age component is lowest in the high-inequality countries of SSA (13 percent), indicating that younger cohorts tend to be as similarly *low* educated as their predecessors. The higher average age contribution in South Asia is mainly driven by the Maldives, reporting about twice (23 percent) the magnitude than India and Bangladesh (see Figure 2 and the discussion above). In contrast, younger cohorts tend to be similarly highly educated as their predecessors in HI OECD countries. MENA and LAC countries show relatively persistent education structures at medium attainment levels, resulting in relatively low contributions of the age component. The relative importance of age is largest in CAE countries (18 percent), again due to the highly dynamic countries in southeastern Europe, as well as in EAP countries (17 percent), where, for example, Hong Kong and Singapore experienced a period of substantial educational expansion.

Generally, the upper-left panel of Figure 3 depicts a negative relation between the relative relevance of the age component and the total education Gini. However, closer inspection of the data to gain information about the underlying process hints at repeating inverted U-shaped relations. First, while the relative contribution of age increased over the whole time span in SA countries, it increased until 1990, but has been decreasing since then, in all other world regions.[17] This implies that the attainment shares of younger cohorts increased relative to their predecessors before they started to stagnate. Second, not only are age-group differences substantially more relevant for females than for males (i.e. $I^E_{\text{ageF}} > I^E_{\text{ageM}}$ in Table 5), but they also exhibit opposing relations with overall educational inequality (see the upper-right panel of Figure 3). The positive relation for males is due to converging education levels of successive male generations, indicating their relative advance in the process of educational improvements. In contrast, the attainment share of young female cohorts is continuously higher than that of their predecessors over the whole time span, resulting in a downward-sloping line. The gender- and time-specific patterns of the age component highlight the varying role of the distribution of education between cohorts in the process of educational expansion, with an increasing divide if younger generations become more highly educated than the older generation, and a narrowing gap as these cohorts age. Beyond that, these findings indicate that before the inclusion of females, educational expansion started among young males.

The contribution of differences between age groups of different sexes becomes increasingly larger than the gender-specific age component as the overall level of educational inequality increases. The divide in educational attainment is thus bigger, comparing, for example, males and females aged 65 and 25, respectively, than

[17]In HI OECD, LAC, SA, and SSA countries, even the absolute level of the age component increased significantly until 1990. These results stem from fixed-effects regressions of each component or its relative contribution over time and are available from the author upon request.

TABLE 5
Summary Statistics: Components, 2010

Region[a]		I^E_{age}	$I^E_{age/sex}$	I^E_{ageF}	I^E_{ageM}	I^E_{sex}	I^E_{within}	$I^E_{residual}$
HI OECD	Level	0.03	0.03	0.02	0.01	0.00	0.03	0.08
	%	15.91	16.08	10.97	4.94	2.61	16.04	49.36
CAE		0.03	0.03	0.02	0.01	0.01	0.03	0.07
		17.44	17.26	12.96	4.47	3.67	16.23	45.39
EAP		0.05	0.05	0.03	0.02	0.01	0.06	0.13
		16.73	17.30	10.48	6.26	3.58	19.22	43.16
LAC		0.04	0.04	0.02	0.02	0.01	0.06	0.15
		13.75	13.41	8.62	5.13	2.61	19.38	50.86
MENA		0.05	0.05	0.03	0.02	0.02	0.09	0.15
		13.71	14.20	8.25	5.46	5.78	24.62	41.70
SSA		0.06	0.08	0.04	0.03	0.05	0.13	0.22
		13.00	15.08	7.63	5.37	8.66	23.08	40.19
SA		0.08	0.09	0.04	0.04	0.05	0.14	0.23
		14.34	15.81	7.33	7.01	8.16	23.09	38.61
Global		0.05	0.05	0.03	0.02	0.02	0.07	0.14
		14.79	15.41	9.50	5.29	4.83	19.84	45.13

[a]CAE, Central Asia and Eastern Europe; EAP, East Asia and the Pacific; HI OECD, high-income OECD; LAC, Latin America and the Caribbean; MENA, Middle East and North Africa; SA, South Asia; SSA, sub-Saharan Africa. For more information on the components, see equation (15) and the explanation that follows.

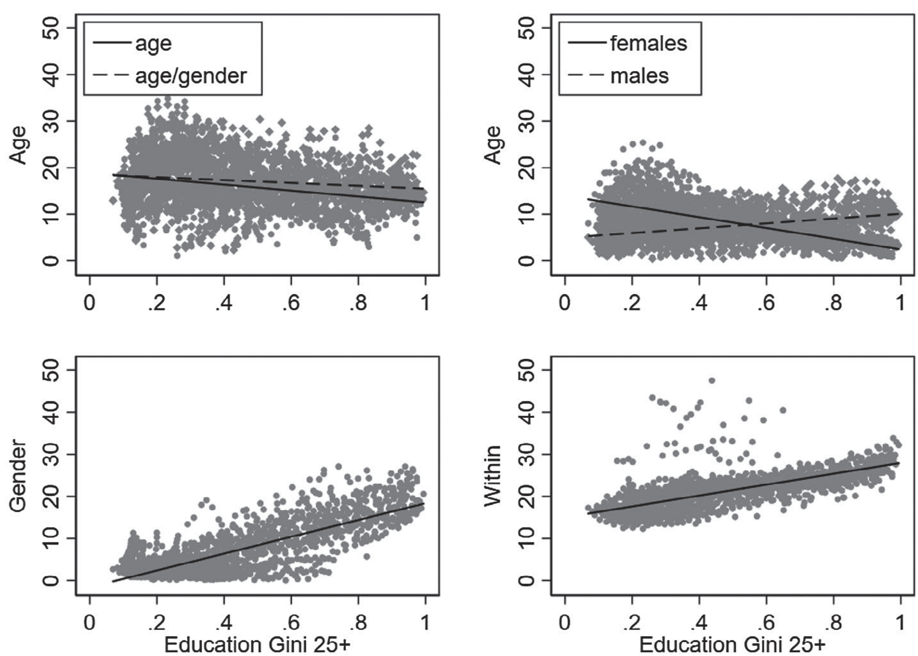

Figure 3. The Contribution of Components (Percent) and Overall Inequality

comparing females of similar age. Figure 4 illustrates this for South Asian countries by plotting age- and gender-specific attainment shares ($\bar{v}_j$) over time. In Nepal, for example, the share of women aged 45–64 in total educational attainment is not only lower than that of younger women, but also significantly lower than that of younger males. This is true for the whole sample period, but of diminishing relevance as the attainment share of young males declines. Conversely, the higher attainment share of young women compared to medium-aged men is relatively low, and even reverses in the 1970s. The upper-left panel of Figure 3 suggests that no significant relation between the age/sex component and overall inequality exists. Yet, disentangling the relation into gender-specific effects in a similar manner as in the upper-right panel of Figure 3 reveals that increasing gaps between young women and older men contribute to lower educational inequality. The reverse is true for gaps between young males and older female cohorts.[18] There is thus an additional factor of gender to age-group inequalities in education.

Closure of the divide between males and females of the same cohorts has significantly contributed to declining educational inequality around the globe. Not only do I observe a positive relation between the total Gini index and the relative contribution of the gender component (see the lower-left panel of Figure 3), but also a significantly decreasing trend of its contribution over time for all world regions. In MENA, SSA, and SA countries, this is due to continuously increasing education shares of young women who follow up in completing basic education levels. In 1970, differences between males and females were almost as relevant in MENA (16 percent) as in SA and SSA countries, both showing a relative contribution of the age component of 17 percent in 1970. But in MENA countries, the decline in the educational gender gap has been more pronounced, and this has enabled the region's stronger reduction in overall educational inequality (see Section 5). Compared to its low overall education Gini, the CAE countries show a relatively high contribution of gender inequality, as gender differences in southeastern Europe have been relatively large. On the other hand, the low value in LAC countries shows that gender differences can be marginal even if overall educational inequality is relatively high. In HI OECD countries, the education shares of both sexes evolve almost simultaneously over the whole sample period, resulting in a constantly low contribution of the gender component. In 82 countries across all world regions, the education share of young females was higher than that of their male counterparts in 2010. This is true for some countries, mainly among the LAC and CAE countries, where young-female shares have been constantly higher than those of young males[19], but mainly for countries where a switch took place in the 1990s. From this, it follows that differential institutions are a relevant factor in shaping gender inequality.

The relevance of inequality between individuals within demographic subgroups of the population is generally decreasing with the level of educational inequality (see the lower-right panel of Figure 3). Nevertheless, inequality within population subgroups makes up the largest share in educational inequality. The relative within-group component contribution is greater than the three between-group components in all regions except CAE. Notably, with a relative contribution of 25 percent in 2010, the within component plays a considerable role in MENA

[18]These results can be obtained from the author upon request.

[19]Argentina, Guadalupe, Jamaica, Martinique, Saint Lucia, Uruguay, Estonia, Latvia, Lithuania, Ireland, Sweden, and Lesotho have higher mean numbers of years of schooling among females than among males throughout the sample period from 1970 to 2010.

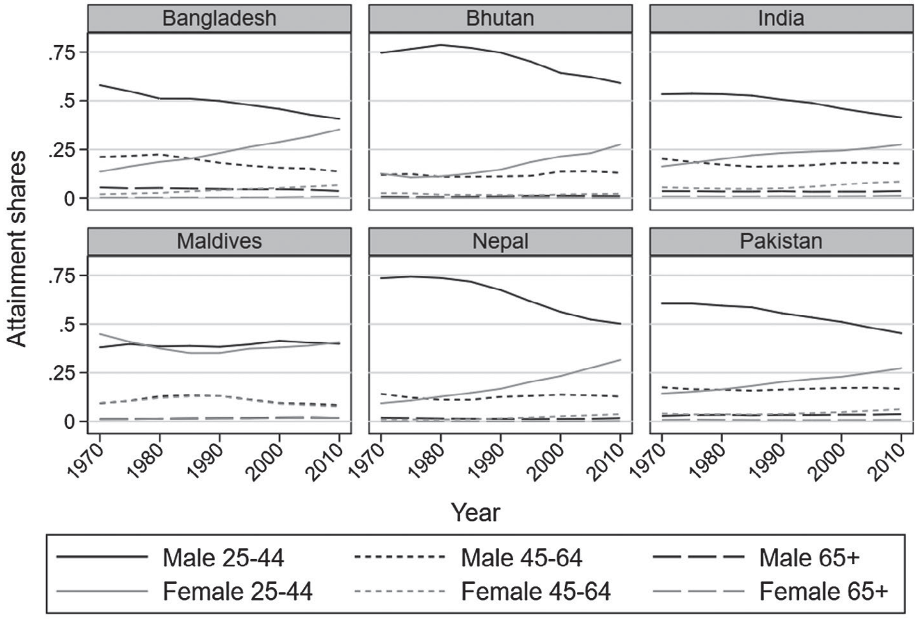

Figure 4. Attainment Shares ($\bar{v}_j$) by Age Group and Gender

countries, compared to the level of educational inequality in the region. This is mainly due to outliers in the Middle East[20], characterized by particularly high within-group inequality. Beyond that, a fixed-effects regression of the within-group contribution over time shows that the relative relevance of within-group inequality is only marginally decreasing over time.[21]

The residual component explains a large part of educational inequality in each world region. While its absolute level decreases with the education Gini, its relative contribution rises as overall educational inequality declines (see the upper panel of Figure 5). In LAC countries, the relatively large residual contribution of 51 percent in 2010 can be explained by its low level of between-group inequality, particularly with regard to gender. In contrast, the residual contribution is low in SSA (39 percent) and SA (40 percent), where differences between males and females are still relevant. Even if this suggests a negative relation of residual inequality with between-group inequality, the lower-left panel of Figure 5 does not provide unambiguous evidence. The residual component approaches zero as gaps between groups vanish. It also tends to be low for high values of between-group inequality. In between the extremes, the variation of the residual component is large. Plotting the residual against the within-group component provides an indication that this is

[20]Quatar and United Arab Emirates.
[21]Globally, the relative contribution of within-group inequality declined by 0.05 percentage points each year. In comparison, the respective time trend for the gender component is 0.13.

due to countries differing with respect to within-group inequality (see Section 7).[22] The residual component thus interacts with the between- as well as the within-group component. First, the lower the between-group inequality, the more subgroups of the population overlap—or, the more permutations are necessary to rank individuals, first, by the average educational attainment of their group and, second, by the relative position within their group. Second, comparing individuals from different groups solely by group averages abstracts from comparisons of outliers not represented by their group's mean attainment. The more polarized groups are—that is, the greater the spread within groups—the more relevant these comparisons become. Hence, the higher the residual component should be.

To sum up, the analyzed components of educational inequality exhibit differential roles across regions and vary over time in the process of educational improvements. Even if the sample period I observe is dominated by equalizing processes between cohorts, differing trends between males and females, as well as over time, indicate that the gap between age groups fluctuates in conjunction with the level of educational attainment. This is also true for differences between sexes. My findings indicate that educational expansion processes at basic levels have predominantly started among young males, before the inclusion of females. However, women outperform men at higher education levels. Nevertheless, closure of the gap between the sexes of equal age groups has significantly contributed to the declining trend of overall educational inequality. The contribution of within-group inequality, in turn, is relatively large and only marginally decreasing over time. Beyond that, polarization within groups impairs between-group comparisons, resulting in larger residual components of educational inequality.

7. The Distributional Impact of Educational Expansion

Convergence of educational attainment within and between age groups as well as sexes has accounted for declining inequality in the distribution of education. However, to what extent has educational expansion around the globe contributed to this trend? Moreover, to what extent have components of educational inequality been affected differently?

Previous work on the distribution of education has found evidence for a strong negative relation between the degree of educational inequality, measured by the education Gini coefficient, and average educational attainment, mostly measured by mean numbers of years of schooling. Yet, Sauer and Zagler (2012) have observed that even if this relation holds across countries, it need not be strictly negative within countries over time. The evidence of Meschi and Scervini (2013) provides support for a U-shaped relation between the average level and the distribution of education. Accordingly, educational inequality has substantially declined as education has expanded, but shifting education structures toward post-secondary levels tends to increase educational inequality. Their findings contrast with the hypothesis of an educational Kuznets curve, which implies an *inverted* U-shape;

[22]This relation is even more visible if I exclude MENA countries with particularly high within-group contributions but low shares of between-group components.

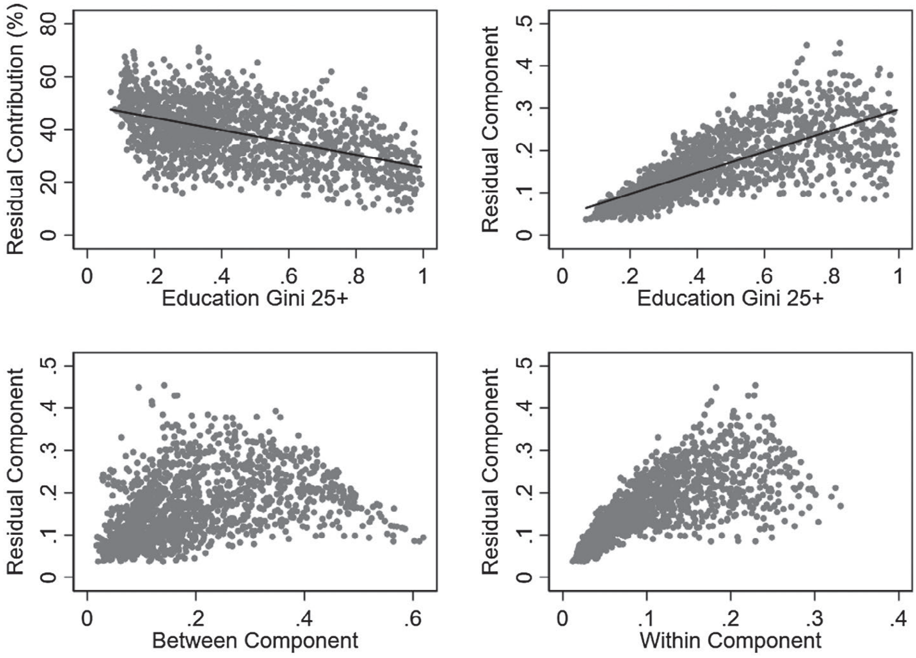

Figure 5. Between-Group, Within-Group, and Residual Components

that is, adverse followed by favorable distributional consequences of educational expansion.

Also, my findings in Section 6 suggest a non-linear relation between educational inequality and average attainment, which is driven by the evolution of between- and within-group inequality. I test whether a non-linear relation between each component and average educational attainment exists in a panel regression framework. Beyond that, I investigate the distributional impact of the education policy to increase compulsory education and test for the presumed negative and positive relations of the residual component with between- and within-group inequality, respectively. A simple model that aims to explain the level as well as the relative contribution of each component can be written as follows:

$$\text{Comp}_{i,t} = \alpha_i + \beta_1 \bar{Y}_{i,t} + \beta_2 \bar{Y}_{i,t}^2 + \beta_3 dur_{i,t}^C + \lambda_t + \varepsilon_{i,t}, \tag{16}$$

where $\text{Comp}_{i,t}$ is the pertinent component of educational inequality or its relative contribution. $\bar{Y}_{i,t}$ is a measure of mean numbers of years of schooling in the population aged 25 and over. Its square accounts for the presumed non-linearity in the relation of interest. $dur_{i,t}^C$ is the formal duration necessary to complete the lower-secondary level, which I use as a proxy for compulsory education. This data is taken from the UIS (see Section 4) and measured as an average over the years during which each of 16 cohorts above age 25 went to school to complete lower-secondary education. Thereby, the time lag between the education policy and educational outcomes is accounted for. Only in the regressions with the residual component as dependent variable, I also include the between- (I_B^E)- and the within-group (I_W^E) components as regressors. I include country- (α_i) and time-specific (λ_t) intercepts. Finally, $\varepsilon_{i,t}$ captures the time-varying component of the error term. I estimate the parameters in equation 16 using a fixed-effects estimator, as the expected value of the time-invariant country-specific effects conditional on the explanatory variables cannot be assumed to be zero.

Tables 6 and 7 present the results using the education Gini as well as the absolute levels and relative contributions of its components as dependent variables. In each case, the squared term of the mean number of years of schooling is significant at the 1 percent level. I thus find strong evidence for the presumed non-linearity in the level–inequality relationship. However, the magnitude and direction of the effects differ between components. Interesting insights can be gained from looking at conditional estimates, computed as $\beta_c = \beta_1 + 2\beta_2 \bar{Y}_{i,t}$. Doing so also enables me to obtain turning points, given by the level of educational attainment at which β_c becomes insignificant before its sign changes.

According to column (1) of Table 6, the total education Gini of the population aged 25 and over is predicted to significantly decrease in the process of educational expansion until 17 years of schooling, but to increase thereafter. The U-shaped relation obtained for both age components in columns (2) and (3) is less pronounced. The part that captures differences between same-sex cohorts (I_{age}^E) becomes insignificant at a level (33 years) well above the sample range of mean numbers of years of schooling. Columns (4) and (5) reveal that the small impact is due to different effects for males and females. While differences between female cohorts increase as education expands up to 3 years, male cohorts converge until 11 years of schooling. This finding provides additional evidence that young males have been the first to benefit from improvements in the transition to universal basic education. In contrast, young women tend to outperform young men at higher education levels. Gaps between males and females are thus predicted to increase as of 8 years of schooling. At the average level of education in South Asia (see Table 2), an increase by 1 year *reduces* the gender component by 31 percent compared to its 2010 level (see Table 5). Conversely, it *increases* the 2010 level of gender inequality in HI OECD countries by a factor of 3.75. Even if the estimated relationship between average educational attainment and within-group inequality equally follows a U shape, the turning point at 13 years of schooling is higher, so that only more recent observations from Canada, Finland, and Germany fall above that level. In contrast, SA countries would be able to reduce their level of within-group inequality by 15 percent in relation to its 2010 level, as average attainment increases by 1 year.

Examination of the relative contributions of the educational inequality components (Table 7) reveals a somewhat different picture. Most importantly, I find evidence for an inverted U-shape of the relation between the relative contribution of both age-related components and mean years of schooling. Thus, as education expands, differentials between age groups become increasingly important, but their relevance for overall inequality vanishes as the educational attainment of successive young cohorts stagnates. Columns (3) and (4) show that the observed educational Kuznets curve with respect to age is driven by the evolution of the distribution

TABLE 6
Regression Results: Components

	(1) I^E_G	(2) I^E_{age}	(3) $I^E_{age/sex}$	(4) I^E_{ageF}	(5) I^E_{ageM}	(6) I^E_{sex}	(7) I^E_{within}	(8) $I^E_{residual}$
$\bar{Y}$	−0.106***	−0.014***	−0.014***	0.004***	−0.018***	−0.031***	−0.030***	−0.049***
	(0.002)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.002)
$\bar{Y}^2$	0.003***	0.000***	0.000***	−0.001***	0.001***	0.002***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
dur^C	−0.063***	−0.018*	0.008	−40.002	−0.016**	0.022**	−0.006	−0.062***
	(0.017)	(0.009)	(0.010)	(0.005)	(0.007)	(0.011)	(0.007)	(0.011)
I^E_B								−0.606***
								(0.013)
I^E_W								0.130***
								(0.046)
Cons.	1.491***	0.296***	0.066	0.051	0.244***	−0.052	0.279***	1.054***
	(0.157)	(0.084)	(0.091)	(0.045)	(0.063)	(0.096)	(0.060)	(0.103)
Obs.	1,539	1,539	1,539	1,539	1,539	1,539	1,539	1,539
R^2	0.942	0.621	0.640	0.459	0.698	0.715	0.881	0.778
N	171	171	171	171	171	171	171	171

Notes Standard errors in parentheses. $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

TABLE 7
Regression Results: Contribution of Components

	(1) I^E_{age}	(2) $I^E_{age/sex}$	(3) I^E_{ageF}	(4) I^E_{ageM}	(5) I^E_{sex}	(6) I^E_{within}	(7) $I^E_{residual}$
$\bar{Y}$	1.339***	2.045***	1.970***	−0.632***	−2.565***	−0.767***	0.000
	(0.163)	(0.161)	(0.097)	(0.096)	(0.126)	(0.084)	(0.000)
$\bar{Y}^2$	−0.064***	−0.071***	−0.090***	0.026***	0.142***	0.009*	−0.000
	(0.009)	(0.009)	(0.006)	(0.006)	(0.007)	(0.005)	(0.000)
dur^C	−0.229	3.105*	0.711	−0.939	4.957***	3.410***	0.000
	(1.770)	(1.757)	(1.059)	(1.050)	(1.366)	(0.912)	(0.000)
I^E_B							−1.000***
							(0.000)
I^E_W							−1.000***
							(0.000)
Cons.	12.651	−18.143	−5.633	18.284*	−27.992**	−6.049	100.000***
	(16.081)	(15.964)	(9.625)	(9.539)	(12.417)	(8.283)	(0.000)
Obs.	1,539	1,539	1,539	1,539	1,539	1,539	1,539
R^2	0.222	0.255	0.478	0.453	0.580	0.285	1.000
N	171	171	171	171	171	171	171

Notes Standard errors in parentheses. $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

of education among females. The relative relevance of gaps between women of different age cohorts increases as education expands up to 8 years of schooling and declines thereafter, while the reverse is true for gaps between male cohorts. In turn, the relative relevance of generational differences between the sexes increases up to 12 years of schooling, but is relatively constant thereafter. The estimated relation between average educational attainment and the relative contribution of gender- and within-group inequality resembles that obtained from looking at the absolute levels of the components. While the turning point is equal for the gender component, it is much later, at 24 years of schooling, for the within-group component. For the latter, the estimated effects in both directions are small, which indicates that the relevance of within-group inequality is relatively constant in the process of educational expansion.

The results in Table 6 further show that increasing the formal duration of compulsory schooling contributes significantly to declining educational inequality via reducing generational gaps between males. However, as they are the first to benefit from the education policy, differences between males and females become larger: Extending the duration of compulsory education significantly increases the absolute level as well as the relative contribution of gender inequality. Moreover, gaps between age groups of different sexes become relatively more relevant. The broader provision of basic education to young males thus contributes to the generation of educational gender gaps. Finally, the absolute level of inequality between individuals within demographic subgroups of the population, and its relative contribution, are positively affected by increasing the duration of compulsory schooling.

The results in the last column of Table 6 provide support for the presumed relations between within-group, between-group, and residual inequality, described in Section 6. The residual component tends to be low if between-group inequality is large and high if education is unequally distributed within population subgroups. Taken together, these findings can explain the high variation of the residual component with increasing between- and within-group inequality, respectively (see the lower panel of Figure 5). Interestingly, the average number of years of schooling and the formal duration of compulsory schooling exert an impact on residual inequality that is additional to their effects via between- and within-group inequality. In contrast, changes in the relative relevance of the residual component are fully explained by changes in the relative contribution of the other Gini components (see Table 7).

8. Robustness

The use of the Gini coefficient in order to measure inequality between and within population subgroups has the disadvantage that it is not additively decomposable. Hence, residual inequality explains up to 50 percent of overall educational inequality. In Section 6, I provide an interpretation of the residual term that relates to the degree of within-group inequality. In other applications (see, e.g., Lakner and Milanovic, 2013), the residual component is thus often added to within-group inequality. Nevertheless, the question remains whether the main results regarding the evolution of educational inequality components and their relation to average educational attainment hold when resorting to the family of decomposable inequality measures (Shorrocks, 1980). To test for the robustness of my results, I thus computed the Theil index to measure educational inequality and decomposed it into age, gender, and within-group components. In Appendix C, I provide information on the computation, as well as figures that compare the two inequality measures (Figures C.1 and C.2) and show the relation between each component's inequality contribution and the overall level of educational inequality as measured by the Theil index (Figure C.3).

The Theil index ranges from 0 to 5 and is positively related to the Gini coefficient. As the Theil index is not bounded from above, the slope rises as the Gini coefficient approaches 1. The levels and relative contributions of the inequality components based on the Theil index provide a similar picture of between- and within-group inequality to that obtained from decomposing the Gini coefficient. However, the relative relevance of inequality between age groups and gender is consistently lower for the Theil as compared to the Gini components. Conversely, within-group inequality is substantially higher, resulting in a global average contribution of 84 percent. This indicates that the residual inequality obtained from decomposing the education Gini is largely absorbed by the within-group component based on the Theil index. Moreover, the within-group component of the Theil index is significantly increasing over time in CAE, MENA, and SSA countries, and constant in the other regions, except for LAC countries. Finally, I am also generally able to resemble the non-linear relations between each component and average educational attainment depicted in Tables 7 and 7 (these results are available from the author upon request).

9. Summary and Conclusions

This paper provides an integrated analysis of the evolution of educational inequality within and between demographic subgroups of the population in the process of educational expansion. For 171 countries over the time span from 1970 to 2010, I decomposed the education Gini coefficient of the total population aged 25 and over into components that measure differences in average educational attainment between age groups, males and females, and a within-group component. Investigation of these new indicators provides insights into educational expansion trajectories around the globe.

Educational inequality between age groups, between males and females, and between individuals from different socioeconomic backgrounds within demographic groups exhibits differential roles across world regions and varies over time as societies become educated. In general, my results provide an indication that educational inequality changes in waves with educational attainment. As long as specific groups are the first to benefit from improvements, inequality first rises but then decreases as larger parts of the population take part in education, and remains constant until further advancements again exert disequalizing effects. Over the sample period that I observe, educational inequality has substantially declined, but tends to increase as education expands further. This U-shaped relation is in line with previous evidence provided by Meschi and Scervini (2013) and contradicts the hypothesis of an educational Kuznets curve.

Each component contributes differently to the overall trend. The relative relevance of divides between age groups increases as younger generations become more highly educated and decrease as these cohorts age. I find evidence on gender-specific cohort effects which indicates that the expansion of basic educational levels has predominantly started among young males, before the inclusion of females. Yet, educational expansion has contributed significantly to closing the gap between the sexes of equal age groups, which has in turn added to the declining trend of overall educational inequality throughout the observed sample period.

In contrast, young women tend to be more highly educated than their male counterparts, particularly in countries with high levels of education. The educational gender gap is thus predicted to increase in the transition to higher education levels. Inequality in the distribution of education within population subgroups accounts for a large part of overall educational inequality and its relevance is only marginally decreasing over time. Thus, the ethnic background, the place of residence, or the social and economic status of people continues to determine the educational prospects of people around the globe. My findings indicate that even if the transition to universal basic education has contributed to declining inequality, shifting of the education structure toward post-secondary levels is not only able to increase gaps between age cohorts, but also between males and females, as well as between individuals within demographic groups. Hence, as long as societies are segmented along various lines, policies benefit particular groups of people. A comprehensive understanding of the distribution of education between and within population groups can help to identify these target groups, thereby helping to improve the distributional impact of policies aimed at educational expansion.

The differential impact of policies on demographic and social groups is especially important in South Asia, where the gender- and within-group differences are among the largest in the world. This is true for supply-side policies that address teaching quality, implement regular monitoring of schools and students' achievement to increase school accountability and empower parents and children, or establish funding equalization schemes and build small schools in order to reduce the urban–rural gap. But demand-side policies such as the elimination of fees—not only for elementary education, but also at higher levels—stipends and conditional cash-transfer programs, as well as information campaigns, should also be tailored to promote particularly vulnerable groups. A prominent example is Bangladesh's girls' stipend program, implemented in 1993. Requirements regarding attendance and scoring in examinations have been combined with other effective measures to increase girls' attainment—for example, providing latrines and safe drinking water—thus helping to substantially increase gender parity in secondary education (Bing, 2009). Following the elaboration of Dréze and Sen (2013) about the "centrality of education," incorporating/anchoring educational policies in a broader commitment by the state that follows universalistic principles in the provision of essential public services is key to reducing educational inequalities.

The new indicators of educational inequality between- and within demographic subgroups of the population enable further research on the causes, policy implications, macroeconomic consequences, and broader societal effects of educational inequality. For example, the age component measures the extent of educational expansion and can be used to examine the consequences for economic growth. Relating the measures of educational gender inequality to democratization, female labor force participation, or fertility can contribute to the understanding of the role of women in development processes. Moreover, the new indicators can be used to analyze whether the closure of education gender gaps has contributed to narrowing gender wage gaps at the aggregate level.

References

Ballarino, G., M. Bratti, A. Filippin, C. Fiorio, M. Leonardi, and F. Scervini, "Increasing Educational Inequalities?" in W. Salverda, B. Nolan, D. Checci, I. Marx, A. McKnight, I. G. Tót, and H. van de Werfhost (eds), *Changing Inequalities in Rich Countries*, Oxford University Press, Oxford, 121–45, 2013.

Barro, R. J. and J. W. Lee, "A New Dataset of Educational Attainment in the World, 1950–2010," *NBER Working Paper 15902*, National Bureau of Economic Research, Cambridge, MA, 2010.

Bauer, R., M. Potancoková, A. Goujon, and K. C. Samir, "Populations for 171 Countries by Age, Sex, and Level of Education Around 2010: Harmonized Estimates of the Baseline Data for the Wittgenstein Centre Projections," *IIASA Interim Report IR-12-016*, 2012.

Benaabdelaali, W., S. Hanchane, and A. Kamal, "Educational Inequality in the World, 1950–2010: Estimates from a New Dataset," in J. A. Bishop and R. Salas (eds), *Inequality, Mobility and Segregation: Essays in Honor of Jacques Silber, Research on Economic Inequality*, Vol. 20, Emerald Publishing, Bingley, U.K., 337–66, 2012.

Bing, W. K., "Education and Inequality in the Developing World," in D. Holsinger and W. Jacob (eds), *Inequality in Education: Comparative and International Perspectives*, Springer, Dordrecht, 86–127, 2009.

Birdsall, N. and J. L. Londono, "Asset Inequality Matters: An Assessment of the World Bank's Approach to Poverty Reduction," *American Economic Review*, 87(2), 32–7, 1997.

Braga, M., D. Checchi, and E. Meschi, "Educational Policies in a Long-Run Perspective," *Economic Policy*, 28(73), 45–100, 2013.

Carnoy, M., "As Higher Education Expands, Is It Contributing to Greater Inequality?," *National Institute Economic Review*, No. 215, 2011.

Castelló, A. and R. Doménech, "Human Capital Inequality and Economic Growth: Some New Evidence," *The Economic Journal*, 1128, C187–200, 2002.

Checchi, D., "Does Educational Achievement Help to Explain Income Inequality?," *Working Paper n.11.2000*, Universitá degli Studi di Milano, 2000.

Checchi, D. and H. G. van de Werfhorst, "Policies, Skills and Earnings: How Educational Inequality Affects Earnings Inequality," *Socio-Economic Review*, 0, 1–24, 2017.

Cohn, E. and J. T. Addison, "The Economic Returns to Lifelong Learning in OECD Countries," *Education Economics*, 6(3), 253–307, 1998.

Crespo-Cuaresma, J., P. Sauer, and K. C. Samir, "Gini Coefficients of Educational Attainment: Age Group Specific Trends in Educational (In)equality," *WWWforEurope Working Paper No. 6*, 2013.

Cruces, G., C. G. Domench, and L. Gasparini, "Inequality in Education," *UNU-WIDER Working Paper No. 2011/93*, 2011.

Dorius, S. F., "The Rise and Fall of Worldwide Education Inequality from 1870 to 2010," *Sociology of Education*, 86(2), 158–73, 2012.

Dréze, J. and A. Sen, *An Uncertain Glory. India and its Contradictions*, Princeton University Press, Princeton, NJ, 2013.

Fan, X., V. Thomas, and Y. Wang, "Measuring Education Inequality: Gini Coefficients of Education," *World Bank Working Paper 2525*, The World Bank, Washington, D.C., 2001.

______, "A New Dataset on Inequality in Education: Gini and Theil Indices of Schooling for 140 Countries, 1960–2000," *Mimeo*, The World Bank, Washington, D.C., 2002.

Goujon, A., K. C. Samir, M. Speringer, B. Barakat, M. Potancooková, J. Eder, E. Striessnig, R. Bauer, and W. Lutz, "A Harmonized Dataset on Global Educational Attainment between 1970 and 2060—An Analytical Window into Recent Trends and Future Prospects in Human Capital Development," *Journal of Demographic Economics*, 82, 315–63, 2016.

Harsha, A., P. Glewwe, and S. Wisniewski, "The Impact of School Quality, Socio-Economic Factors and Child Health on Students' Academic Performance: Evidence from Sri Lankan Primary Schools," *Education Economics*, 21, 2–37, 2013.

Inter-American Development Bank, *Facing up to Inequality in Latin America: Economic and Social Progress in Latin America, 1998–99 Report*, Johns Hopkins University Press, Baltimore, MD, 1999.

Kremer, M., N. Chaudhury, F. H. Rogers, K. Muralidharan, and J. Hammer, "Teacher Absence in India: A Snapshot," *Journal of the European Economic Association*, 3(2/3), 658–67, 2005.

Lakner, C. and B. Milanovic, "Global Income Distribution: From the Fall of the Berlin Wall to the Great Recession," *World Bank Policy Research Working Paper 6719*, The World Bank, Washington, D.C., 2013.

Lam, D. and D. Levison, "Declining Inequality in Schooling in Brazil and its Effects on Inequality in Earnings," *Journal of Development Economics*, 37(1–2), 199–225, 1991.

Londono, J. L., "Kuznetsian Tales with Attention to Human Capital," *Paper presented at the Third Inter-American Seminar in Economics*, Rio de Janeiro, Brazil, 1990.
López, R., V. Thomas, and Y. Wang, "Addressing the Education Puzzle: The Distribution of Education and Economic Reform," *World Bank Working Paper 2031*, The World Bank, Washington, D.C., 1998.
Maas, J. and C. Criel, "Distribution of Primary School Enrollments in Eastern Africa," *Wold Bank Staff Working Paper 511*, The World Bank, Washington, D.C., 1982.
Meschi, E. and F. Scervini, "Expansion of Schooling and Educational Inequality in Europe: The Educational Kuznets Curve Revisited," *Oxford Economic Papers*, 66, 660–80, 2013.
______, "A New Dataset On Educational Inequality," *Empirical Economics*, 47, 695–716, 2014.
Morrisson, C. and F. Murtin, "The Kuznets Curve of Human Capital Inequality: 1870–2010," *Journal of Economic Inequality*, 11(3), 238–301, 2013.
Potancoková, M., K. C. Samir, and A. Goujon, "Global Estimates of Mean Years of Schlooling: A New Methodology," *IIASA Interim Report IR-14-005*, 2013.
Ram, R., "Educational Expansion and Schooling Inequality: International Evidence and Some Implications," *The Review of Economics and Statistics*, 72(2), 266–74, 1990.
Rama, M., T. Béteille, Y. Li, P. K. Mitra, and J. L. Newman, *Addressing Inequality in South Asia*The World Bank, Washington, D.C., 2015.
Rosthal, R. A., *Measures of Disparity: A Note, Research report*, Killalea Associates, Inc., 1978.
Sauer, P. and M. Zagler, "Economic Growth and the Quantity and Distribution of Education: A Survey," *Journal of Economic Surveys*, 26(5), 933–51, 2012.
______, "(In)equality in Education and Economic Development," *Review of Income and Wealth*, 60, 353–79, 2014.
Schofer, E. and J. W. Meyer, "The Worldwide Expansion of Higher Education in the Twenthieth Century," *American Sociological Association*, 70(6), 898–920, 2005.
Schuetze, H. G., "International Concepts and Agendas of Lifelong Learning," *Compare*, 36(3), 289–306, 2006.
Sheret, M., "Equality Trends and Comparisons for the Education System of Papua New Guinea," *Studies in Educational Evaluation*, 14(1), 91–112, 1988.
Shorrocks, A. F., "The Class of Additively Decomposable Inequality Measures," *Econometrica*, 48, 613–25, 1980.
Silber, J., "Factor Components, Population Subgroups and the Computation of the Gini Index of Inequality," *The Review of Economics and Statistics*, 71(1), 107–15, 1989.
Swainson, N., *Girls' Education in South Asia*, Oxfam, Oxford, 2006.
World Bank, *Student Learning in South Asia: Challenges, Opportunities, and Policy Priorities*, The World Bank, Washington, D.C., 2013.

Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's web site:

Appendix A: Between-Group Inequality
Appendix B: The Sample of Countries by World Region
Table B.1: Regions & Countries Appendix C: Theil Index of Educational Inequality
Figure C.1: Education Gini & Theil Index, Population Aged 25+
Figure C.2: Comparison: Contribution of Components (%)
Figure C.3: Contribution of Components (%) & Overall Inequality (Theil)