- Permanyer, I.; Boertien, D. (2018) "A century of change in global educational inequality and gender differences in education". Presented in Session 12: Global Trends in Educational Stratification. *Population Association of America (PAA). 2018 Annual Meeting.* Denver (United States), April 26-28. https://paa.confex.com/paa/2018/meetingapp.cgi/Paper/19987

A century of change in global educational inequality and gender differences in education

Iñaki Permanyer[1] and Diederik Boertien

Abstract: This article aims to explore how overall education inequality is articulated in four basic constituent parts (inequality among women and men, inequality favoring men and inequality favoring women) and how these parts have jointly evolved over time in the world and its regions. We investigate whether or not these regions have gone through periods in which, as education expands, both education inequality and the gender gap in education declined simultaneously. For that purpose, we have developed a new inequality index adapted to the ordinal nature of educational attainment with nice decomposability properties that can be useful for scholars and policy-makers to track the evolution of overall education inequality and all its subcomponents. Our findings suggest that (i) the composition of education inequality has been shifting dramatically over time; in particular (ii) inequality among men was usually higher than inequality among women until the turn of the century; from then onwards women's and men's educational distributions are equally unequal; (iii) the educational advantage of women over men has become the most important source contributing to educational inequality for most high- and middle-income countries, and (iv) the world and its regions have gone (and are expected to go) through considerably long periods of time where education expansion is not at odds with both gender and education inequality reductions.

[1]Corresponding author. Centre d'Estudis Demogràfics, Carrer de Ca n'Altayó, Edifici E-2, Campus de la UAB, 08193 Cerdanyola del Vallès, Spain. E-mail: inaki.permanyer@uab.es. The research leading to these results has received funding from the European Research Council (ERC-2014-StG-637768, EQUALIZE project), the CERCA Programme from Generalitat de Catalunya and from the Spanish Ministry of Economy and Competitiveness "Ramón y Cajal" Research Grant Program (RYC-2013-14196).

1. Introduction

For a long time, a global expansion of mass schooling has been sweeping the globe. This process, which has been observed in all countries around the world, has been highly beneficial for an ever growing number of individuals, so this is a worthy cause for celebration (Morrisson and Murtin 2009). While a growing body of scholarship has shed considerable light on the 'efficiency part' of the process (i.e. the country-level average education attainments are reasonably well-documented), much less is known about the 'equality part'. Whether or not some countries or certain groups of individuals within those countries benefit from education expansion disproportionately more than others is a matter of great scientific interest and policy concern which, unfortunately, has not received much scholarly attention and is still quite poorly understood.

One of the facets of education inequality that has indeed attracted the attention of many scholars is the gender gap in education. During the last decades we are witnessing the emergence of a "global first" phenomenon: the closing and reversal of the gender gap in educational attainment. For the first time in history, in many high- and middle-income countries, young women are attaining higher levels of education than young men, and the same trend is expected to occur in other countries as well in the coming decades (Esteve et al 2016). While relevant and informative, the comparisons of average attainments for women and men typically ignore the potential heterogeneity that might exist within those groups. Yet, implicitly treating women and men as if they were homogeneous groups does not allow addressing key questions about the education distribution. For instance: Is the education distribution among women becoming more unequal as their average attainment surpasses that of men? Is the education inequality among women higher than that of men? Does the reduction – and eventual reversal – of the gender gap in education go hand-in-hand with less educational inequality overall? How are these inequalities related with one another? While it seems clear that (i) overall education inequality, (ii) education inequality among women and among men, and (iii) education inequality favoring women and favoring men are interrelated with one another, no previous study has attempted to investigate and flesh out the nature of that relationship. For the first time, in this paper we put together these different ingredients into a coherent whole to explore how overall education inequality is decomposed and relates to its basic constituent parts. Such decomposition is extremely useful for scholars and policy-makers to go beyond purely descriptive results and analyze what factors are the most important drivers of inequality and its evolution over time.

The importance of knowing the trends of these different inequalities stems from their differential impacts on societies around the world, which are multiple and might go in opposite directions. On the one hand, increasing education inequality – either for the population as a whole or for women and men separately – 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 and general well-being – see Dorius 2013, Ballarino et al 2014). On the other hand, the closing and reversal of the gender gap in education in favor of women might increase female labor force participation (particularly in better-paying

formal-sector jobs), prolong the duration of women in the labor force (which in turn could lead to more efficient economic outcomes; Klasen and Lamanna 2009) and generate more egalitarian attitudes (both in the domestic and public spheres, Esteve et al 2016) that could eventually shift the balance of power towards females (Grant and Behrman 2010). Symmetrically, high levels of gender inequality in favor of men have long been considered to be detrimental for the economic performance of countries and the well-being of their inhabitants (Klasen and Lamanna 2009). Given the benefits of both high levels of overall and gender equality in education, it becomes relevant to ask whether both types of equality can be maximized simultaneously, and in which situations possible trade-offs between both objectives can emerge. These are key questions to which we aim to give answers in this article.

What do we know about global trends in education inequality and the gender gap in education? As regards the former, opposing findings from different studies – partly driven by the use of alternative data sets, methods and indicators – have not facilitated the creation of a consensus among scholars. At most, a majority of researchers could agree that education inequality has followed an inverted U-shaped trajectory over time, whose maximal height and turning point depend on the geographical region we are dealing with (Morrisson and Murtin 2009, 2013, Dorius 2013) – an admittedly simplistic vignette. As regards the gender gap in education, most studies suggest that its evolution over time has not been monotonic either (Wils and Goujon 1998, Grant and Behrman 2010). While in its initial stages the expansion of education predominantly benefited men, in later stages women caught-up, thus reducing and even reversing the gender gap. What this sketchy account does not provide is a precise description of the *timing* in which both phenomena take place. Depending on whether the trend reversal of education inequality occurs after or before the gender gap turnaround, countries might experience periods in which trends in both types of inequality run in opposite directions, or in which both types of inequalities increase or decrease at the same time. If it turns out that at some point they start declining simultaneously, an opportunity window opens up where educational expansion brings the additional benefits of reducing both overall and gender inequality in education (i.e. a period where there is no dilemma between 'efficiency' and 'equality'). Whether or not the world and its regions have gone (or can still expect to go) through such a period will be a main empirical goal of our analysis.

Summing up, this article aims to explore how overall education inequality is articulated in its different constituent parts and how these parts have jointly evolved over time in the world and its regions to identify if, and for how long, they have reaped the benefits of education expansion alongside simultaneous reductions in overall inequality and its components. To attain our research objectives, we have created a new measure of education inequality specifically crafted to simultaneously meet two requirements that are not met by currently existing inequality measures. First, our measure decomposes overall inequality in two components: inequality within groups (including inequality among women and among men) and between groups (which can be further decomposed in inequalities in favor of women and inequalities in favor of men). Second, it has been designed to handle the ordinal data we are dealing with in this paper (currently, virtually all education inequality measures are based on cardinal information,

like ‘years of schooling’). Hopefully, the new measure proposed in this paper can be a useful addition to those practitioners aiming at gauging inequality in the context of ordinal variables.

The empirical analysis relies on the latest version of the Barro and Lee (2015) database (henceforth BL) for the period from 1950 to 2010 and on some of our own projections, extending the results up to 2040. The huge geographic coverage of the database (146 countries) allows performing both global and regional analyses over time. Our findings suggest that (i) the composition of education inequality has been shifting dramatically over time; in particular (ii) inequality among men was usually higher than inequality among women until year 2000 approximately, from then onwards women’s and men’s educational distributions are equally unequal; (iii) the educational advantage of women over men has become the most important source contributing to educational inequality for most high- and middle-income countries, and (iv) educational expansion initially implied increased overall education inequality and (for a shorter period of time) higher gender inequality, but most world regions have gone or will go through periods where such expansion is not at odds with gender or overall education inequality reductions.

2. Background and hypotheses

It is difficult to overestimate the importance of education for the well-being of individuals across their entire life cycle. In general, highly educated individuals tend to have a better command over resources, higher levels of employment and better-paid jobs, they have longer and less unequal lifespans, lower risk of getting divorced, of being poor or materially deprived and higher levels of subjective well-being (Farkas, 2003; Härkönen and Dronkers, 2006; Lafortune, 2013; Meara et al., 2008; Psacharopoulos & Patrinos, 2004). Indeed, education is a key stratification variable in demographic behavior (see Lutz, Butz and KC 2014). For these reasons, the education expansion that has been sweeping the world during the last decades is to be highly welcomed. This expansion includes rising literacy rates (Crafts 2002) and increases in schooling enrollment rates and in completed years of primary, secondary and college education (Benavot and Riddle 1988, Benavot et al. 1991; Meyer, Ramirez, and Soysal 1992; Ramirez and Meyer 1980; Barro and Lee 2000; Cohen and Soto 2007; Morrisson and Murtin 2009). Regarding college education, by 1970, 6.4% of the world’s population aged 25-29 had obtained a college degree. Three decades later, this proportion had increased to 13%, and the expected figure for 2050 is 29.4% (KC et al. 2010).

Despite its relevance for many socio-demographic and economic outcomes – e.g. high levels of education inequality are associated with higher economic inequality and unemployment levels, lower social cohesion and higher poverty rates (Ballarino et al 2014) – researchers have only recently started to study how individuals’ education is distributed across and within countries (e.g. Castelló and Domenech 2002, Benaabdelaali et al. 2012, Dorius 2013, Meschi and Scervini 2013, Morrisson and Murtin 2009, 2013 or Jordá and Alonso 2016). While these studies differ in many respects (e.g. they typically employ different indicators, datasets and methodological approaches, and/or their geographic and temporal coverage does not

necessarily coincide) they all cohere with the following narrative. As education started to expand at some point back in time during the modern era – predominantly in favor of men – education inequality began to increase. After several decades, as education expansion gradually shifted the population from low to highly educated categories, education inequality reached its highest point and afterwards started declining until the present day[2]. In other words: empirical evidence seems to support the existence of a Kuznets curve in education. This highly simplified narrative generally applies for the world as a whole, its different regions and most of its countries.

A notable feature of the education expansion process is that it has not been gender neutral (Wils and Goujon 1998, Grant and Behrman 2010, Dorius and Firebaugh 2010). Despite initially favoring males, the gender gap has closed rapidly in recent years and, in many countries, even reversed in favor of women (Esteve, Garcia and Permanyer 2012), a trend that is expected to continue over the next decades (KC et al. 2010; Lutz and KC, 2011). In 1970, men represented 63.6% of the total college educated population. This percentage decreased to 52.6% in 2000 and it is likely to reach 44% in 2050, with most high-income countries reaching lower levels (KC et al 2010). The recent and prospective sex-specific share shifts across the education ladder – usually in favor of women – have attracted the attention of several researchers because of their potentially sizeable impacts across the board (e.g. fertility changes, higher female labor force participation and higher economic growth, or shifting power relations both in public and private domains; see Esteve et al 2012, 2016, Van Bavel 2012, Van Bavel and Klesment 2017). It should be pointed out that the gender gap in education – which is currently favoring women in most high- and middle-income countries – is not as large as it used to be when it favored men in the initial stages of the education expansion.

Joint inequality trends: alternative hypotheses and scenarios

Lacking any empirical evidence we can rely upon, we cannot but hypothesize what might be the relationship between overall education inequality (henceforth denoted as I) and its 'basic subcomponents': inequality among men (I_m), inequality among women (I_f), and the two subcomponents of the gender gap in education (i.e. inequality favoring men (A_m) and inequality favoring women (A_f)). A priori, there are many possible ways in which these subcomponents might have evolved over time in a way that fits the aforementioned stylized narratives on the trends in overall education inequality and the gender gap in education. In this regard, Figure 1 plots what might have been the hypothetical trajectories of I, I_m and I_f over time. In line with the existing empirical evidence, overall education inequality is posited to first increase and then decrease. What about the education inequality among men and among women? Since the education expansion initially benefited men, we expect I_m to increase earlier than I_f. Yet, since the education expansion for women took place in a shorter period of time,

[2] The studies by Castelló and Domenech 2002, Benaabdelaali et al. 2012, Meschi and Scervini 2013 and Jordá and Alonso 2016 focus on the second half of the 20th century, so they cover the downward part of the inverted U-shape trajectory. The studies by Dorius 2013 and Morrisson and Murtin 2009, 2013 use data from the 19th to the 20th century, so they cover the upward *and* downward portions of the trajectory.

we expect I_f to increase faster than I_m. Based on the evidence reporting the closing and reversal of the gender gap in education in favor of women we expect I_f to exceed I_m at some point in time (denoted as t in Figure 1). Yet, given that it is unlikely that obstacles to access education, which women experienced in the past, will be put in place for men, we do not expect the differences in these two quantities to be as large as they might have been during the first stages of the education expansion that predominantly benefited men (i.e. we expect the two curves to remain relatively close to each other after time t). Whether or not this has actually been the case for the different world regions will be investigated in the empirical section of the paper.

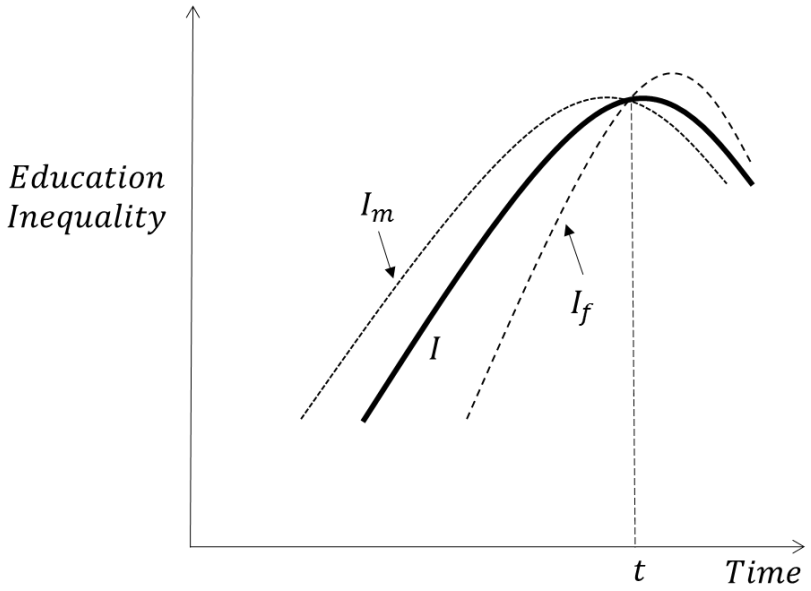

Figure 1. Schematic relationship between overall education inequality (I) and inequality among women (I_f) and men (I_m). Source: Authors' elaboration.

What about the relationship between trends in overall education inequality (I) and the gender gap in education? Are they expected to move in the same or in opposite directions? For ease of exposition, the latter will be denoted as G (an indicator taking negative and positive values, with negative values indicating gender gaps in favor of men, positive ones in favor of women, and 0 indicating equality between women and men; details given below). In Figure 2 we show a non-exhaustive sample of hypothetically plausible trajectories for I and G that are consistent with what current empirical evidence tells us about their non-monotonic trends (see above). In these diagrams, t_0 denotes the time when G reaches its minimum, t_G the time when the gender gap goes from negative to positive (i.e. when $G = 0$), and t_I the time when overall inequality I reaches its maximum. The only substantive difference between the scenarios shown in Figure 2 is the relative position of t_0, t_I and t_G vis-à-vis each other and the time that has elapsed between them (i.e. the *timing* of these events).

Take the first scenario, where the gender gap reversal occurs before overall education inequality reaches its maximum (i.e. $t_G < t_I$). For brevity, it will be referred to as an 'early reversal scenario' (see panel A in Figure 2). In this case, between years t_0 and t_G, we observe increases in education inequality in tandem with reductions in the gender gap in education (i.e.

inequalities running in opposite directions). Afterwards, between years t_G and t_I we observe simultaneous increases in education inequality *and* in the gender gap in education. Consider now the second scenario, where the gender gap reversal occurs after overall education inequality reaches its maximum (i.e. $t_G > t_I$). Symmetrically, this will be referred to as a 'delayed reversal' scenario (see panel B in Figure 2). Here, between years t_0 and t_I we observe increases in education inequality but reductions in the gender gap. Afterwards, between years t_I and t_G, there are reductions in education inequality *and* the gender gap in education. Lastly, panel C in Figure 2 depicts a hypothetical scenario in which overall education inequality and the gender gap in education reach its maximum the same year (i.e. $t_0 = t_I$). For obvious reasons, it will be referred to as a 'synchronic scenario'[3]. In that case, one would observe that overall education inequality and the gender gap in education would always go in the same direction. For the years preceding t_0 both inequalities would increase simultaneously, and the opposite would happen after that year.

In *all* these scenarios, there are periods in which overall education and gender inequality either go (i) in opposite directions, (ii) in the same normatively undesirable direction (i.e. simultaneously increasing), or (iii) in the same normatively desirable direction (i.e. simultaneously decreasing). In case (i), there are trade-offs among inequality types, i.e. one type of inequality is reduced at the expense of the other. In (ii) there is an important efficiency-equality dilemma: education expansion is restricted to some groups of individuals and countries in such a way that both education inequality and the gender gap in education tend to increase over time. Lastly, in (iii) we observe the opposite trend: education expansion occurs in such a way that both education inequality and the gender gap in education reduce over time (i.e. there is no efficiency-equality dilemma). When this happens, we say that countries are passing through an opportunity window to potentially reap the benefits that such inequality reductions entail.

How should we define the starting and ending points of such opportunity window? While the starting point is clearly set in the year where both overall education and gender inequality start declining simultaneously, the choice of the ending point can be more controversial. The difficulty stems from the fact that, since small gender gaps in favor of women are not necessarily undesirable from a normative point of view[4], one might want to extend the

[3] It is important to highlight that the terms 'early', 'delayed' or 'synchronic' are merely descriptive and have no normative connotation, they just aim at describing the chronology of different events.

[4] While there is a broad consensus that gender gaps favoring men should count as 'gender inequality', the reverse situation generates some controversy. On the one hand, gender gaps in education favoring women are often not considered to be normatively undesirable. This point of view can be justified on the grounds that women's edge in education somehow compensates the disadvantage they experience in other domains (e.g. lower wages or labor force participation). In this line, Grant and Behrman (2010) suggest that women's educational advantage might be explained because "in contexts where labor markets and cultural traditions are gender biased, young women may be aware that education might be the only opportunity to increase their status in society", thereby implying that female educational advantage should not be really counted as 'gender inequality'. In this line, the Gender Gap Index (GGI) from the World Economic Forum caps the gender gaps at one, tacitly assuming that gender gaps in favor of women are de facto equivalent to 'gender equality'. On the other hand, some composite measures of gender disparities count *any* departure from complete equality as 'gender inequality', no matter if it favors women or men (e.g. the United Nations' Gender Development Index (GDI) and Gender Inequality Index (GII) or the Gender Equality Index (GEI) from the European Institute for Gender Equality). In this paper, we adopt an

opportunity window period until the gender gap goes beyond a certain upper bound $U \geq 0$ – as long as overall education inequality I continues to decline. The year in which such threshold is attained will be denoted as t_U (it is illustrated in panel B of Figure 2). Clearly, when $U = 0, t_U = t_G$.

One of the main aims of this paper is to report the different time periods in which world regions either experience (i) trade-offs among inequality types, (ii) increases in both types of inequalities, or (iii) simultaneous decreases (i.e. report if they experience the aforementioned efficiency-equality dilemmas or not).

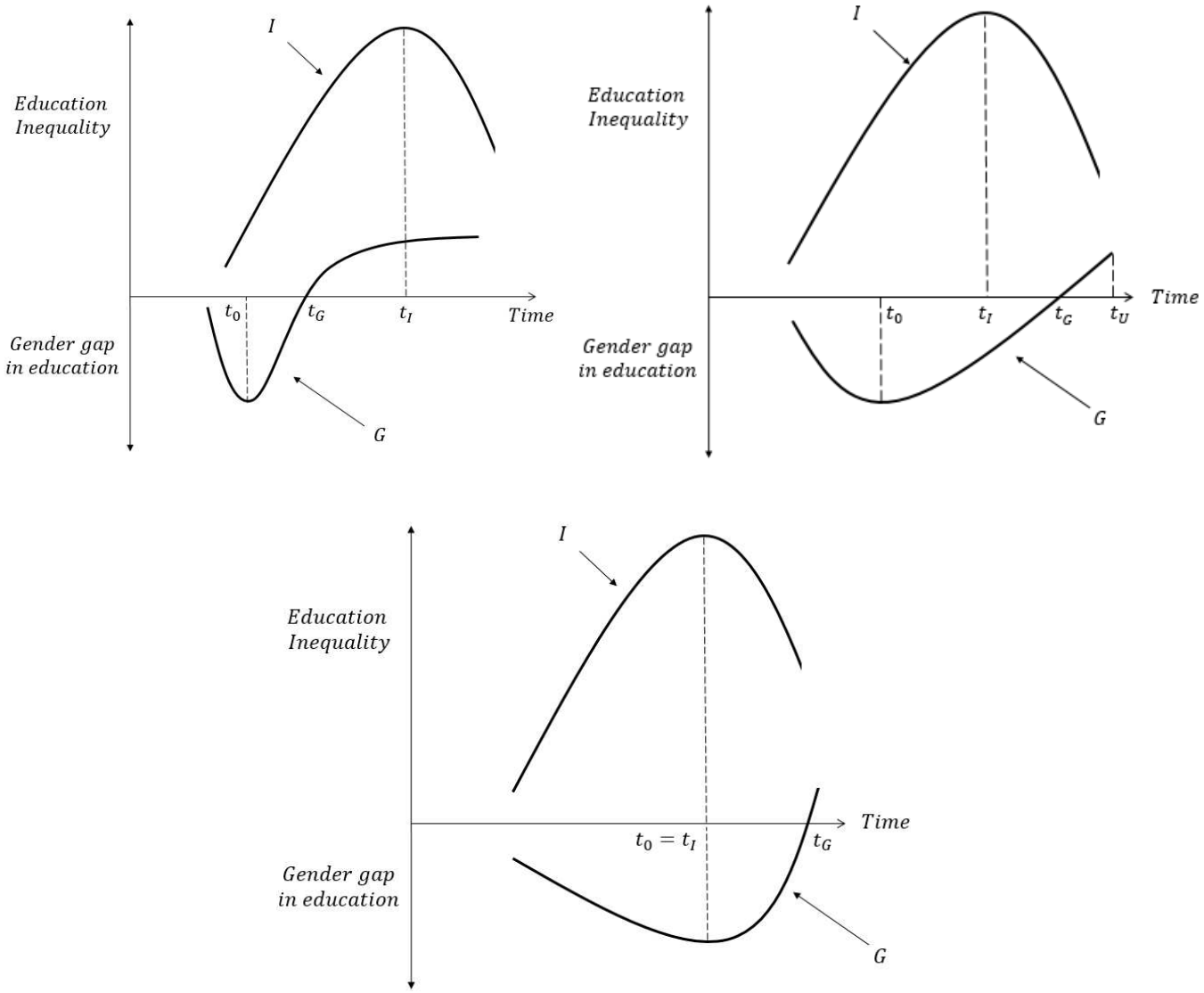

Figure 2. The timing of education inequality and the gender gap in education in different hypothetical scenarios: (A) Early reversal scenario: $t_G < t_I$, (B) Delayed reversal scenario: $t_G > t_I$, and (C) Synchronic scenario: $t_0 = t_I$. Source: Authors' elaboration.

intermediate position and consider that sizeable departures from complete gender equality beyond a loosely defined threshold are normatively undesirable.

3. Data

Like most studies on education inequality we have used the dataset of Barro and Lee (2015), which is currently the largest in terms of geographic and temporal coverage. The data is based on a compilation of censuses and surveys by UNESCO and provides comparable information on 7 educational attainment categories (no education, primary uncompleted, primary completed, secondary uncompleted, secondary completed, tertiary uncompleted and tertiary completed) for 146 countries during the period 1950-2010. Whenever appropriate, these categories can be reduced to 4 (no education, some primary, some secondary and some tertiary).

We used the estimates from Barro and Lee on educational attainment for the period 1950-2010 without further adjustments. At the same time, we decided to calculate our own projections of future educational attainment using a similar procedure to Barro and Lee's. Calculating our own predictions allowed us to focus on the age group 30-34, instead of the age group 15-64 for which the projections of Barro and Lee were calculated. Selecting this age group enables us to minimize the influence of changes in the age structure within countries on our estimates and at the same time will show the most recent changes in educational attainment (by the age of 30-34, the vast majority of individuals have achieved their highest educational attainment). Another motivation has been that some implausible values were observed in the Barro and Lee projections.[5]

Projections

Our projections of future educational attainment are based on a logistic growth curve model estimated for the period 1950-2010 that allows for country-specific time trends (random slopes). The model, estimated separately for men, women, and each of the three educational stages, can be expressed as follows:

$$ln\left(\frac{s_{jt}}{100 - s_{jt}}\right) = \alpha_j + \beta_j t + \mu_{jt} \qquad [1]$$

where s_{jt} is the share of the population having attended educational stage j in year t. The coefficient β_j in this case estimates the (linear) time trend in the (logistically transformed) share of the population attending the given educational stage. The model estimates country-specific constants (α_j) as well as separate coefficients β_j for each country (random slopes in a multi-level model where the two levels are countries and years, μ_{jt} constitutes the error term), so that the predicted educational expansion over time is allowed to follow different trajectories across countries. In this model, countries are expected to eventually approach a 100% attendance for primary and secondary education. It appears unrealistic to expect countries to

[5]In some Western countries the attainment of tertiary education for the age group 25-64 was predicted to decline dramatically with time (e.g. Australia, Austria, Finland).

converge to a 100% attendance of tertiary education. We therefore arbitrarily set a ceiling of 70% for tertiary education. As shown later, predictions using this ceiling fit the data well.

The coefficients from these models are used to project attendance of educational stages for the period 2015-2040. In order to arrive at the 7 educational categories of the historical data, we multiplied attendance shares with the completion rates observed in 2010 for each given country and gender (we multiplied the share of the population that is predicted to attend an educational stage in year *t*, by 'share completed in 2010' / 'share attended in 2010').[6] To safeguard comparability across historical time periods and projections, we decided to present predicted results based on our growth models for all the time period 1950-2040. In the results section we examine the fit of our predicted numbers to actual numbers, which appears to be highly accurate.

4. Methods

In this paper we treat educational attainment as an ordinal variable. This is a non-trivial decision with important implications that stands in sharp contrast to previous studies on education inequality across individuals[7]. There are several reasons why we have made this decision. On the one hand, cardinal variables like 'years of schooling' can be a poor proxy of the substantive type of education individuals might have received and their interpretation can be biased because of (i) the country-specific duration of different education cycles, and (ii) the existence of repeaters. In addition, 'years of schooling' is quite prone to measurement error due to recall bias. On the other hand, ordinal variables are much less prone to measurement error. As opposed to what happens with 'years of schooling', the meaning of the ordinal variable categories (e.g. attaining primary, secondary or tertiary education) is reasonably comparable across countries and over time.

The major disadvantage of using ordinal variables is that their variability cannot be ascertained with well-known inequality measures, like the Gini coefficient or the Theil index among many others (indeed this is *the* key practical reason why education inequality studies have always relied on cardinal variables). One of the contributions of this paper is to partially fill this gap enlarging the practitioner's toolkit by proposing a new inequality measure specifically designed for ordinal variables.

A new measure of education inequality for ordinal data

The tools available to assess inequality are substantially reduced when working with ordinal variables (the main reason being that the notion of 'mean' – which is crucial in the definition

[6] In a small set of cases, predicted levels of attending secondary education exceeded predicted levels of primary education attainment by a small margin. To assure that the shares of educational categories eventually summed up to 1, in those cases we set attendance of primary education to the level of secondary education attendance.

[7] Studies like Castelló and Domenech (2002), Benaabdelaali et al. (2012) or Jordá and Alonso (2016) cardinalize the Barro and Lee dataset using different techniques (the first two estimate the average length of each education cycle while the last one fits a continuous distribution to capture 'within-cycle variations'). Others, like Meschi and Scervini (2013) or Morrisson and Murtin (2009, 2013), work with the cardinal variable 'years of schooling'.

of cardinal inequality measures – is not meaningful in the ordinal case; see Allison and Foster 2004). In this section we propose a new measure of inequality for ordinal variables that has useful decomposability properties. This measure indicates the probability that two randomly chosen individuals have *different* education attainments. Let k be the generic number of education categories we will be working with (in our case $k = 7$) and let N be the population size. We will denote the number of individuals in the population with educational attainment i (with $1 \leq i \leq k$) as N_i, and $p_i = N_i/N$ will be the corresponding population share. We define our ordinal inequality measure as

$$I(p_1, \cdots, p_k) \coloneqq \sum_{i=1}^{i=k} \sum_{j=1}^{j=k} p_i p_j \mathbb{I}(i,j) \qquad [2]$$

where $\mathbb{I}(i,j)$ is an indicator function that takes a value of 1 whenever $i \neq j$ and 0 otherwise. Whenever all individuals have the same educational achievement (i.e. $p_i = 1$ for some category i) there is no inequality, so $I = 0$. For any other distribution, I takes strictly positive values. The index is maximized whenever the population is evenly spread out across all education categories[8] (i.e. $p_i = 1/k$ for all $i = 1, \cdots, k$).

A useful feature of the ordinal inequality index suggested here is that it is nicely decomposable when the population is partitioned in different groups. In this paper we consider the partition of the population between women and men (N^f and N^m denote their corresponding population sizes), but any other population partition in an arbitrary large number of groups would do as well. Let N_i^f, N_i^m be the number of women and men with educational attainment i (hence $N_i^f + N_i^m = N_i$). Their relative shares are denoted as p_i^f, p_i^m. It is straightforward to check that our inequality index can be decomposed as

$$I = s_f I_f + s_m I_m + s_b \left(A_f + A_m\right) \qquad [3]$$

where

$$I_f = I\left(p_1^f, \cdots, p_k^f\right) \qquad [4]$$

$$I_m = I(p_1^m, \cdots, p_k^m) \qquad [5]$$

$$A_f = \sum_{i=2}^{i=k} \sum_{j<i} p_i^f p_j^m \qquad [6]$$

[8] A priori, one might expect inequality to be maximized when the population is split in two equally sized groups located at the extremes of the distribution (formally, when $p_1 = p_k = 1/2$ and all other $p_i = 0$). Even if it is not entirely clear whether such distribution maximizes inequality or polarization (see Esteban and Ray 1994), such inequality measures exist and have already been proposed in the literature (see Abul Naga and Yalcin 2008 and Kobus and Milos 2012). Unfortunately, these measures are *not* amenable to the additive decomposition we propose in this paper, and yet their values are highly correlated with the ones given by our measure (see details below).

$$A_m = \sum_{i=2}^{i=k} \sum_{j<i} p_i^m p_j^f \qquad [7]$$

$$s_f = N^f(N^f - 1)/N(N-1) \qquad [8]$$

$$s_m = N^m(N^m - 1)/N(N-1) \qquad [9]$$

$$s_b = 2N^f N^m/N(N-1) \qquad [10]$$

The derivation of equation [3] is shown in Appendix 1. I_f and I_m are the inequalities within the groups of women and men respectively, A_f, A_m the inequalities between women and men favoring women and favoring men respectively, and $s_f, s_{m,} s_b$ represent the relative weight of each component depending on the population size of each group. The between-group component A_f (resp. A_m) measures the probability that a randomly selected woman has higher (resp. lower) educational attainment than a randomly selected man. Therefore, A_f and A_m can be interpreted as women's educational advantage over men and vice versa. Using these two components we define the extent of gender (in)equality in a given society as $G := A_f - A_m$, a measure ranging from -1 to 1. If $G = -1$, there is no woman whose educational attainment is higher than that of any man, and if $G = 1$, the reverse is true. When the education distributions of women and men are identical, $G = 0$. This measure is not monotonic: values above 0 reflect a better state of affairs for women and vice-versa. The additive decomposition formula shown in [3] articulates into a coherent whole overall education inequality and its four basic subcomponents.

Currently there are only a few other measures of inequality adapted to the ordinal setting (see Abul Naga and Yalcin 2008 and Kobus and Milos 2012). Unfortunately, such measures do not admit the additive decomposition shown in equation [3], so they have not been considered in this paper[9]. It turns out that our index for ordinal variables is intimately related with two other classical measures of heterogeneity defined in the context of nominal and cardinal variables respectively: the 'index of fractionalization' and the 'Gini coefficient'. Indeed, all three measures have much in common because they are grounded in the same basic principle: two individuals are picked at random and one inspects whether (or to what extent) they share a given characteristic/attribute. The only difference between the nominal, ordinal and cardinal cases is the metric that is used to assess the similarity between pairs of individuals. In the nominal case one inspects whether the two individuals belong to different pre-specified groups or not, in the ordinal one whether one of the corresponding attainments is superior to the other or not and in the cardinal case one takes into account the distance between the corresponding

[9] For the sake of completeness we have compared our results with the ones we would obtain using the alternative inequality measures proposed by Abul Naga and Yalcin (2008) and Kobus (2015). It turns out that they are highly correlated (r=0.86) – details not shown here but available upon request.

(cardinal) attainments. Not surprisingly, all three measures have an extremely similar functional form[10].

Interpreting inequality decompositions

Equation [3] is reminiscent of well-known additive decompositions of inequality in within- and between-group components for the cardinal case. Yet, the interpretations in both cases are entirely different. In the cardinal case, the between-group component is the inequality that would be observed in a hypothetical distribution where each individual had the same educational attainment as the *mean* in his group. Yet, it does not make sense to speak about the contribution of between-group inequality to overall inequality in the ordinal context because the notion of 'mean' is not applicable. Instead, the term $s_b(A_f + A_m)$ in equation [3] should be interpreted as the proportion of pairwise comparisons where individuals have different educational attainments involving a woman and a man. To illustrate the difference between both approaches consider a hypothetical scenario where the number of women and men and their educational attainments turned out to be exactly the same. For the cardinal case the between-group component would go to zero because the gender-specific means would be the same. In the ordinal setting proposed here, the between-group component would amount to 50% because half of the education comparisons between pairs of randomly selected individuals would involve a woman and a man. Indeed, in such hypothetical scenario the contribution to overall inequality of the four subcomponents shown in [3] would be exactly the same: 25%. These fundamental differences should be taken into account when interpreting the results.

5. Empirical findings

In this section we present our empirical findings based on the systematic exploration of the 7 education categories' distributions that are reported in the Barro and Lee (2015) dataset. We start documenting the global education expansion and the gender gap reversal during the period 1950-2040 using our model predicted data (see section 3). Afterwards, we calculate the corresponding education inequality index together with its four subcomponents, and report their values and trends – both globally and regionally – for the same time periods. Finally, we analyze the relationship between overall education inequality, its components and the gender gap in education.

Education expansion

Figure 3 presents the share of women having attended primary, secondary, and tertiary education for 7 world regions as well as the predicted shares based on our models. Whereas

[10] The fractionalization index for nominal variables can be written exactly as in equation [2] and the Gini coefficient can be written as $\sum_{i=1}^{i=k}\sum_{j=1}^{j=k} p_i p_j \left|(y_i/\mu)-(y_j/\mu)\right|$, where y_i is the value of the cardinal variable one is interested in for group i and μ is the mean of the distribution. In this respect, our index *I* can be thought as the 'missing link' between the index of fractionalization and the Gini coefficient.

attendance to primary education was only close to being universal in Advanced Economies and Eastern Europe in 1950, shares close to universality have now reached East Asia and Latin America, and this pattern is expected to extend to South Asia, the Middle East and Africa by 2040. Attendance to secondary education was still a minority phenomenon in all world regions in 1950, but today a majority of individuals in most world regions attends secondary education. In all regions but South Asia and Sub-Saharan Africa more than 90% of women are expected to attend secondary education by 2040. Tertiary education was not very extended in the 1950s, but almost half of the women aged 30-34 are expected to attain tertiary education by 2040 in all regions but Sub-Saharan Africa, where tertiary education is predicted to increase only slowly. Figures for men are similar and displayed in Appendix 2.

For the period 1950-2010, Figure 3 shows both predicted and actual shares of attendance. The fact that the two sets of shares are hardly distinguishable from each other in the graph reflects the good fit of the logistic growth curve models described in section 3. Only for secondary education in Eastern Europe and primary education in East Asia some differences between both become visible in certain time periods. Due to this relatively neat fit between predictions and reality, we believe that presenting our model-predicted results does not imply a qualitatively important loss of information.

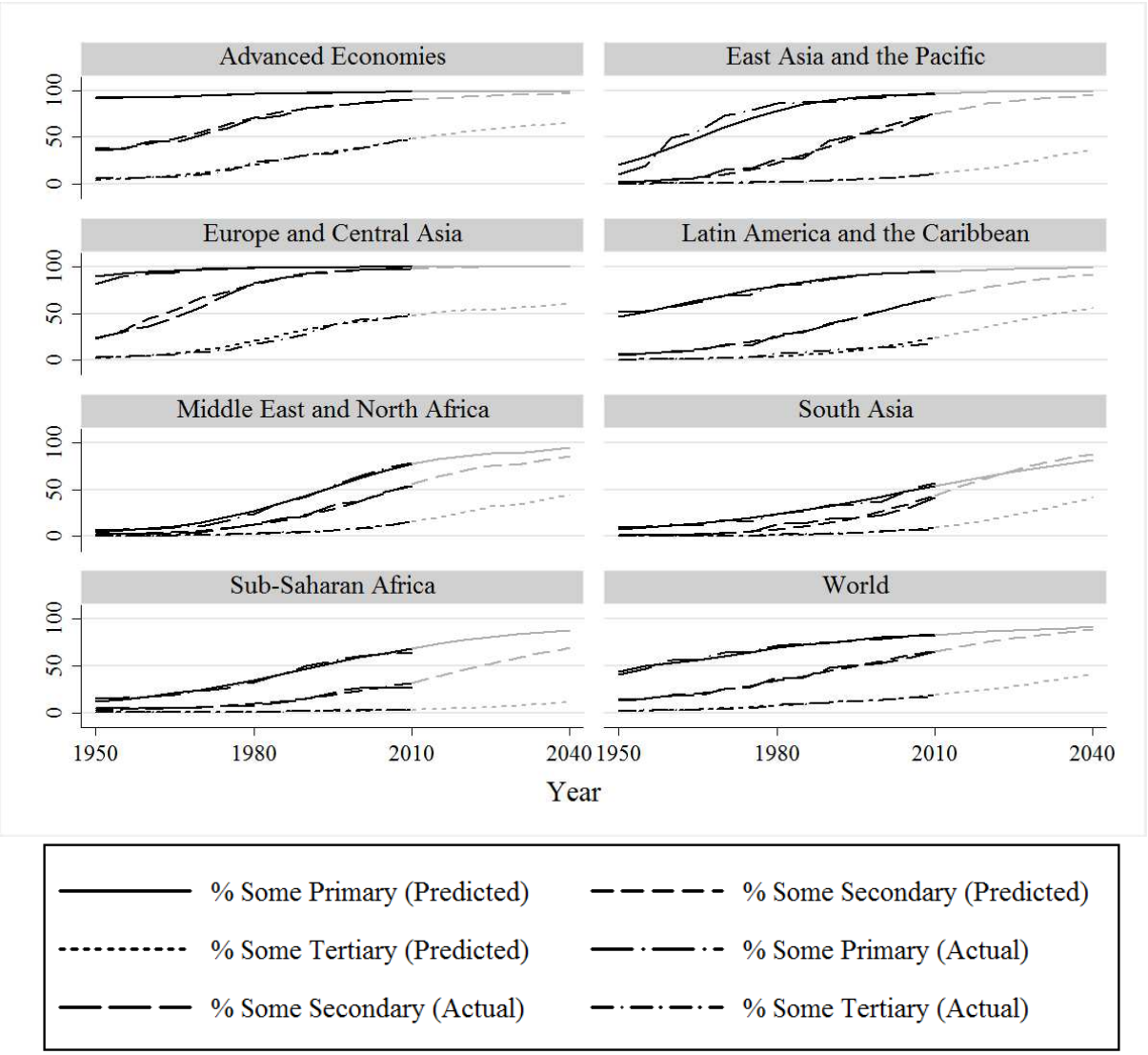

Figure 3. Women's attendance of educational stages 1950-2040, by region, predicted and actual (weighted by population size of countries). Parts in greyscale are predicted values. Source: Authors' calculations based on the BL dataset.

Gender gap reversal in education

We continue to describe the development of educational attainment over time by presenting our measure of the gender gap in education. Figure 4 displays, for the 7 world regions as well as for the world overall, how the gender gap in education has actually developed from 1950 to 2010 and how it is expected to develop across time until 2040. Recall that negative values represent distributions where women are likely to be less educated than men and vice versa. In line with previous research, one can observe that women used to be less educated than men in the past in all world regions, but this gender gap has reduced dramatically across the globe, and even reversed in the Advanced Economies, Eastern Europe, and Latin America well before 2010. The gender gap in these three regions must have reached its historical minimum in a period preceding 1950. The evolution of the gender gap has not always been monotonic: in the

regions of Sub-Saharan Africa, the Middle East and North Africa, East Asia Pacific and South Asia it has initially *declined*, attained its minimum somewhere between 1960 and 1990 (see the first column of Table 2) and then started increasing towards gender equality. By 2040, the gender gap is expected to be closed and slightly reversed in almost all world regions (the only exception being Sub-Saharan Africa). While at the beginning of our observation window we can clearly distinguish two clusters of regions (on the one hand the Advanced Economies, Eastern Europe, and Latin America – henceforth referred to as 'forerunning regions' – and on the other hand Sub-Saharan Africa, the Middle East and North Africa, East Asia Pacific and South Asia – referred to as 'laggard regions') by 2040 most gender gaps are expected to converge towards a slightly positive value indicating some educational advantage in favor of women. Inspecting the 'global gender gap in education' (i.e. taking the entire world as a unit of analysis), we also observe a U-shaped trajectory over time. In 2010, the global gender gap was still in favor of men but this is not expected to be the case anymore by 2040. This story differs from the conclusion of an ever-decreasing gap in education between men and women over time once taking a relative measure of educational differences (e.g. Barro and Lee, 2015, Chapter 2.6).

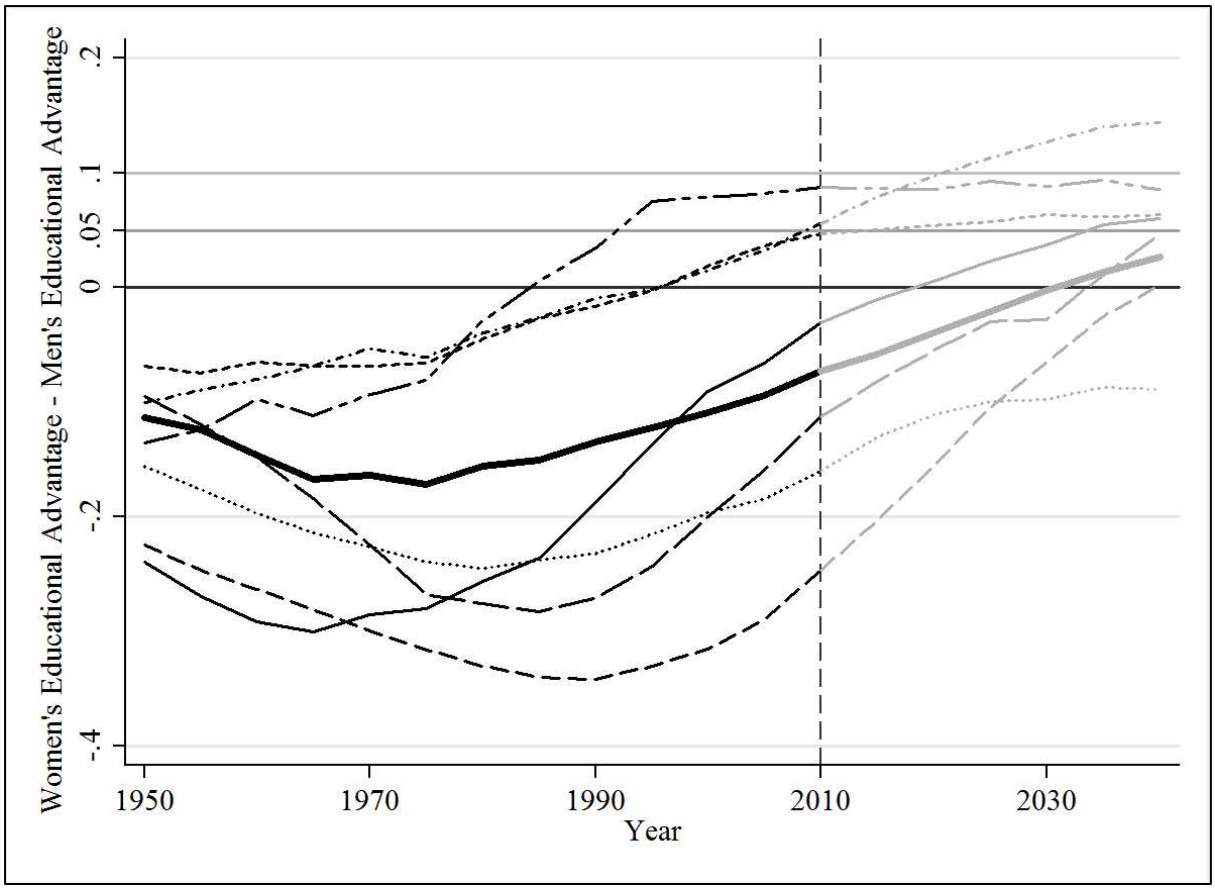

Figure 4. Gender Gap in Education across the period 1950-2040, weighted by country population size. Parts in greyscale are predicted values. Source: Authors' calculations based on the BL dataset.

Evolution of Overall Education Inequality and its components

Figure 5 displays the development of education inequality across the world and within world regions based on ordinal educational categories. It also simultaneously displays the four components of the inequality decomposition formula shown in equation [3]. The thick solid line represents overall inequality *I* over time and roughly follows an inverted U-shaped curve worldwide, congruent with the results of Dorius (2013) and Morrisson and Murtin (2013). Each region individually appears to follow a part of that curve during the period under study. In some regions inequality used to be still relatively low in the 1950s, but inequality increased with time and educational expansion. In other regions, inequality was already high but has recently started to decrease slightly. This decrease is expected to continue in the future, and to find its expression in a decline in educational inequality worldwide.

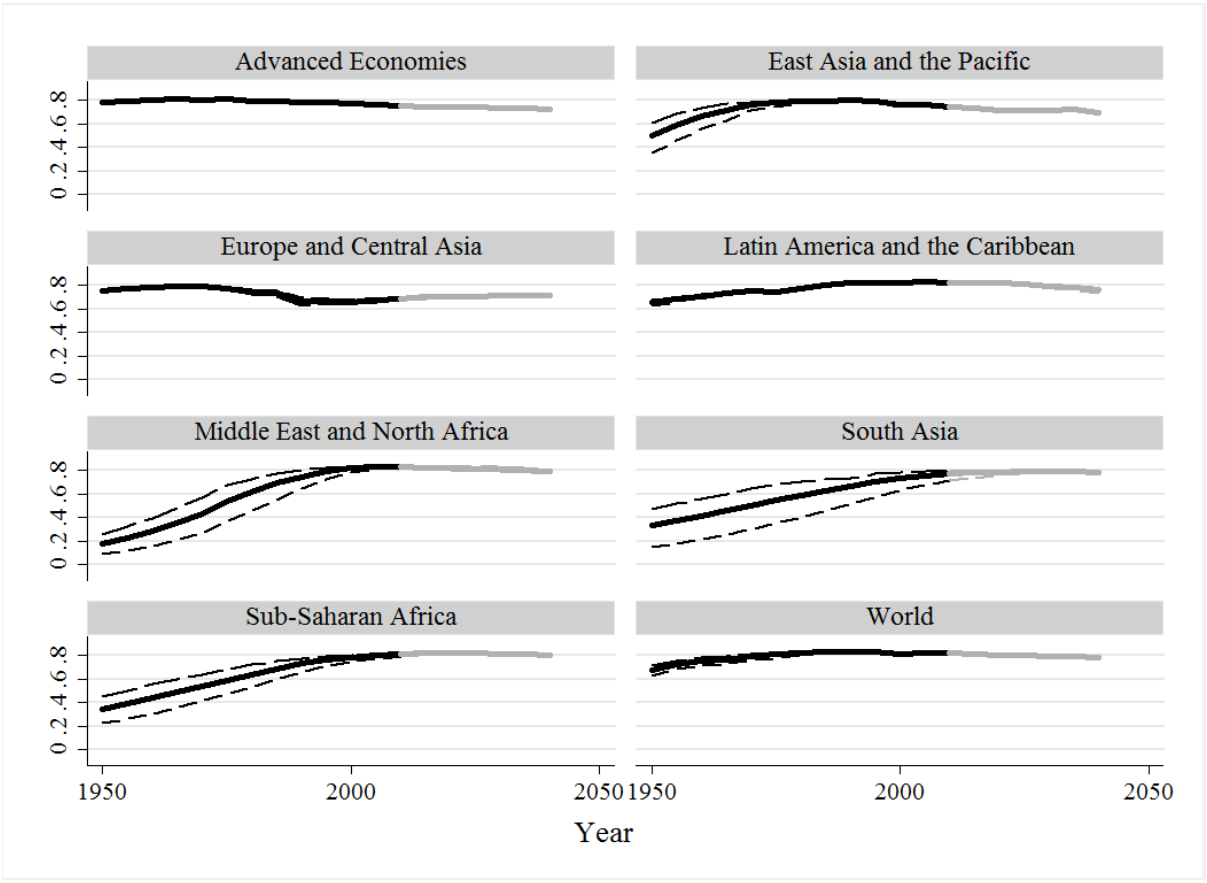

Figure 5. Educational Inequality and its components across time and space, 1950-2040, weighted by country population size. Parts in greyscale are predicted values. Source: Authors' calculations based on the BL dataset.

The other curves of Figure 5 represent educational inequality within women and men (I_f and I_m), as well as the probability that a randomly selected woman has higher education than a randomly selected man (A_f) and vice versa (A_m). The lines for inequality within men and women are only visible during the period 1950-2000 and only for those regions with lower levels of education (i.e. the laggard regions of Sub-Saharan Africa, the Middle East and North Africa, East Asia Pacific and South Asia). In those regions inequality in terms of education among women was much lower than inequality among men in the 1950s. By the 2010s, inequality among men and women became very similar to overall education inequality, so the three curves are indistinguishable from each other in the right tails of the graphs. This means that in the past women used to be a considerably more homogeneous group in terms of education compared to men. Surprisingly, despite the reversal in the gender gap observed and its expected persistence in the future, the groups of men and women are expected to remain equally unequal in the period 2010-2040. The educational disadvantage of men therefore does not seem to translate into them being a more homogeneous group in terms of education compared to women – as we were originally expecting. Overall, this suggests that the stylized trajectories shown in Figure 1 are a rough approximation of what has actually happened with overall and gender-specific education inequalities in the different world regions.

Contribution of the Different Components to Overall Education Inequality

Table 1 displays the relative contribution of each of the four components discussed in the previous section to overall education inequality (these are derived from decomposition formula [3]). Recall that, as a reference, in a hypothetical country where the educational attainment of men was identical to that of women, the percent contribution of the four components shown in Table 1 would be exactly the same: 25%. As can be seen, the composition of education inequality has been shifting dramatically over time. Back in the 1950s, inequality among women contributed very little to overall education inequality in the laggard regions (e.g. a mere contribution of 9.9% in South Asia), while the opposite was true for inequality among men (e.g. 38.9% in the same region). At the beginning of our observation period, the educational advantage of men over women was by far the main contributor to education inequality in the world as a whole and in most of its regions (see last column in Table 1). Sixty years later, that contribution has decreased substantially at a global level. Indeed, in the forerunning regions (i.e. in most high- and middle-income countries) it has become the least important contributor to education inequality (i.e. 23.5%, 21.9% and 23.3% for the advanced economies, Eastern Europe and Latin America in 2010). Concomitant with these changes, the educational advantage of women over men has become an increasingly important ingredient of overall education inequality (see second-to-last column in Table 1). Sixty years ago, the contribution of that kind of inequality to overall education inequality was by far the *least* important among the four (particularly so in the poorest regions of the world, e.g. only 9% in South Asia). Nowadays, women's educational advantage is the main contributor to education inequality in

most high- and middle-income countries (i.e. 26.6%, 28.3% and 26.7% for the advanced economies, Eastern Europe and Latin America in 2010).

Region	Year	%CI_f	%CI_m	%CA_f	%CA_m
Advanced Economies	1950	26.8	23.2	22.8	27.2
Advanced Economies	1980	24.5	25.4	23.6	26.5
Advanced Economies	2010	24.4	25.6	26.6	23.5
Advanced Economies	2040	23.7	26.2	27.2	22.8
East Asia and the Pacific	1950	16.1	32.9	13.5	37.5
East Asia and the Pacific	1980	22.8	26.6	17.2	33.4
East Asia and the Pacific	2010	24.2	25.7	24.0	26.1
East Asia and the Pacific	2040	21.7	28.2	27.2	22.8
E. Europe & Central Asia	1950	33.1	17.8	20.1	28.9
E. Europe & Central Asia	1980	25.6	24.3	24.1	26.0
E. Europe & Central Asia	2010	25.8	24.0	28.3	21.9
E. Europe & Central Asia	2040	23.8	26.1	28.1	22.1
Latin America & Carib.	1950	23.7	26.0	21.3	29.0
Latin America & Carib.	1980	25.1	24.8	23.7	26.3
Latin America & Carib.	2010	25.7	24.2	26.7	23.3
Latin America & Carib.	2040	23.4	26.0	30.0	20.6
Middle E.& North Africa	1950	12.1	37.4	11.7	38.7
Middle E.& North Africa	1980	18.6	29.6	14.6	37.2
Middle E.& North Africa	2010	23.2	26.5	21.7	28.5
Middle E.& North Africa	2040	22.9	26.7	26.6	23.7
South Asia	1950	9.9	38.9	9.0	42.2
South Asia	1980	16.0	31.8	12.0	40.2
South Asia	2010	21.7	27.4	17.5	33.5
South Asia	2040	22.3	27.7	25.0	24.9
Sub-Saharan Africa	1950	16.6	31.9	14.4	37.1
Sub-Saharan Africa	1980	21.2	27.4	16.1	35.3
Sub-Saharan Africa	2010	23.7	26.0	20.2	30.1
Sub-Saharan Africa	2040	24.7	25.2	22.3	27.8
World	1950	23.2	26.5	20.9	29.4
World	1980	23.6	26.1	20.4	29.9
World	2010	24.4	25.5	22.8	27.3
World	2040	23.1	26.9	25.9	24.1

Table 1. Inequality in educational attainment by region and year for 25-29 year-olds, decomposed into within and between women/men components. I: Overall inequality; I_f, I_m: Inequality among women and men respectively; A_f / A_m: Probability that a randomly selected woman/man is more educated than a randomly selected man/woman; %C: Percent contribution of the different components. Authors' calculations based on Barro and Lee (2013).

The relationship between gender equality and overall inequality

What can we say about the relationship between the gender gap in education and overall education inequality in the midst of the aforementioned compositional changes? Have they moved in the same or in opposite directions? To give a precise account of the joint levels and trends of I and G, in Figure 6 we present a full-fledged description of how these two indicators have co-varied over time for the world and its regions (the values of I are shown in the vertical axis and those of G in the horizontal one). Putting together the world regions' experiences shown in that figure, we conclude this section presenting a broad-strokes account of the joint evolution of education and gender inequality, which consists in four stages. Stage I: The male-dominated education expansion brings increases in education inequality and drives the gender gap in education in favor of men. In this case, both inequalities temporarily go in the normatively undesirable direction, and might therefore be a period where educational expansion has most externalities. Stage II: The delayed incorporation of women into mass education tilts the gender gap towards the opposite direction and brings further increases in education inequality until it reaches its maximum. Here, both kinds of inequality go in opposite directions and trade-offs between both policy objectives emerge. During this stage potential efforts to slow down education expansion in order to prevent overall education inequality from rising further could be at odds with reducing the gender gap in education. Stage III: When further education expansion gradually shifts the population towards the higher educational categories, education inequality starts declining and the gender gap goes to zero: this is the onset of an opportunity window where education expands while both inequalities decrease simultaneously. Stage IV: Further education expansion, particularly in favor of women, reverses the gender gap in education while further decreasing overall education inequality.

Even if not all regions (let alone the individual countries) fit this stylized description, it reasonably represents the different trajectories shown in Figure 7. For the laggard regions of the Middle East and North Africa, East Asia Pacific and South Asia we can observe the occurrence of stages I, II and III. Until 2010 Sub-Saharan Africa had only completed stages I and II, and it is expected that by 2040 it will have entered in the third one. As regards the advanced economies and Eastern Europe, we observe the occurrence of stages II, III and IV during our 1950-2040 study period. Lastly, the region of Latin America stands out as being the only one that does not fit very well our 4-stage narrative. The initial reductions of the gender gap go in tandem with increases in education inequality and the posterior decreases in education inequality occur when the gender gap increases in favor of women. Whereas future educational expansion in other regions is expected to simultaneously reduce both forms of inequality, expansion in Latin America is likely to increase the gender gap favoring women to relatively high levels.

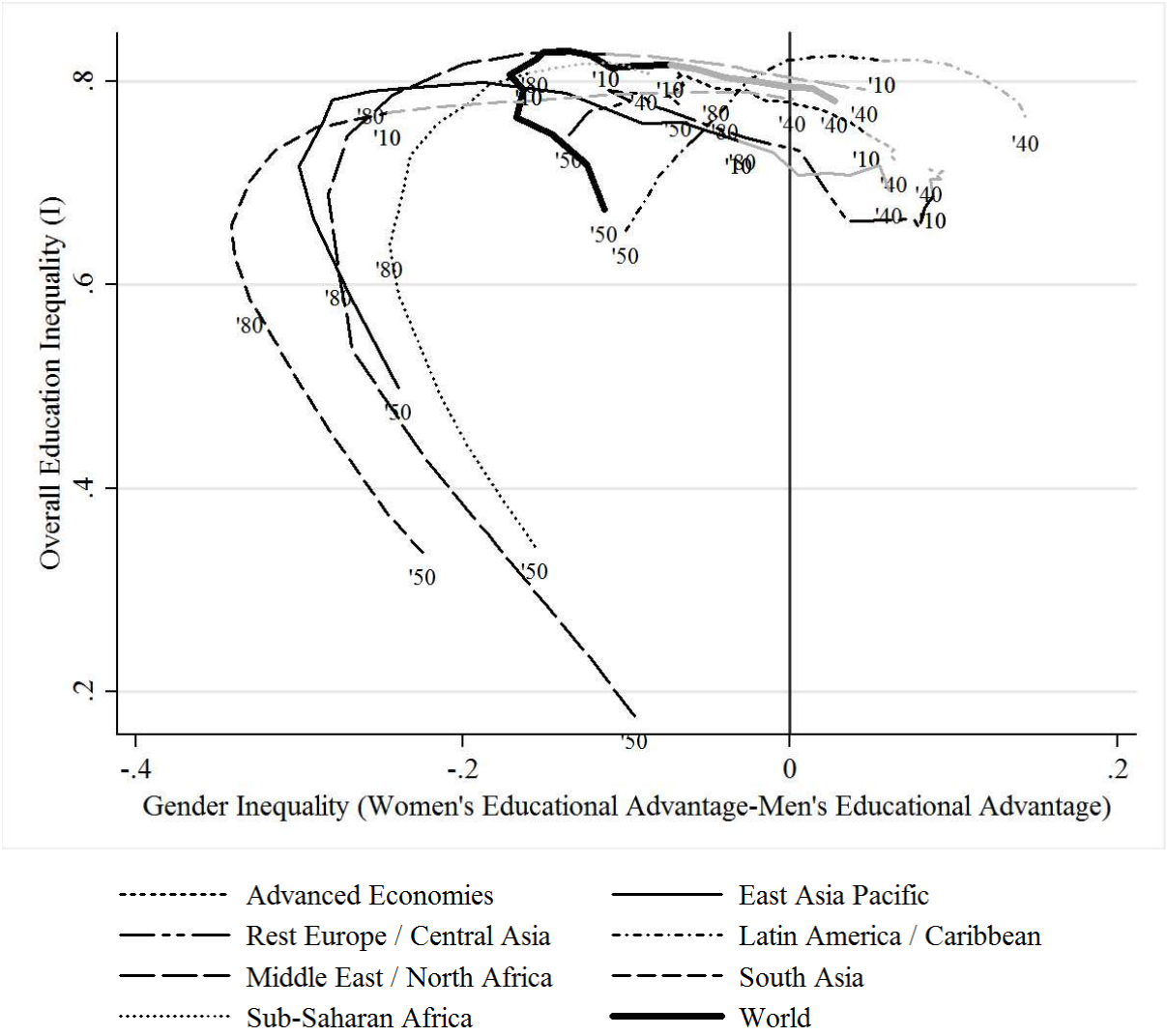

Figure 6. Development of gender and overall inequality in education over time. Parts in greyscale are predicted values. Source: Authors' calculations based on the BL dataset.

Drawing from the results shown in Figure 6, in Figure 7 we indicate the starting and the end of the time periods in which the world and its regions have either experienced (i) trade-offs among these inequality types, (ii) increases in both types of inequalities, or (iii) simultaneous decreases[11]. During the time frame considered here all world regions (except in Latin America) have gone through periods where education expansion has not been an egalitarian process, as it simultaneously increased both kinds of inequalities. Generally, this period has been longer for the laggard regions (e.g. 40 years in South Asia) and considerably shorter for the forerunning ones. Analogously, all regions have experienced considerable periods of time with trade-offs between overall education inequality and the gender gap in education (i.e. periods in which one inequality could only be reduced at the expense of increasing the other). At one extreme, the advanced economies have only experienced such periods for relatively short time spells (15 years), while at the other extreme, Latin America has experienced them almost during the entire time frame (80 years). Overall, there have been many more periods in which

[11] To generate this table, we have assumed that values of the gender gap between 0 and 0.1 (i.e. slight education advantage in favor of women) are *not* normatively undesirable.

education inequality increased in tandem with gender gap reductions than periods with the opposite combination. Lastly, the world and all its regions have also gone through considerably long periods without experiencing the efficiency-equality dilemma. The longer the duration of those periods, the more likely the corresponding countries can reap the benefits associated to the reduction of both types of inequality. While some regions are expected to take advantage of such opportunity window only for a decade (e.g. South Asia, between 2030 and 2040), others have benefited and are expected to benefit from it during 65 years (e.g. advanced economies, between 1975 and 2040). The world taken as a whole entered such period in 1990, and is expected to continue until 2040 with some small interruptions.

Inspecting Figures 2, 4, 5 and 6, we can say that the world as a whole and most of its regions – all except Latin America – have experimented or are likely to experiment in the near future the 'delayed recovery scenario'. The Latin American experience seems to fit the 'early recovery scenario'. Finally, no region has gone through the 'synchronic scenario'.

[[[Figure 6]]]

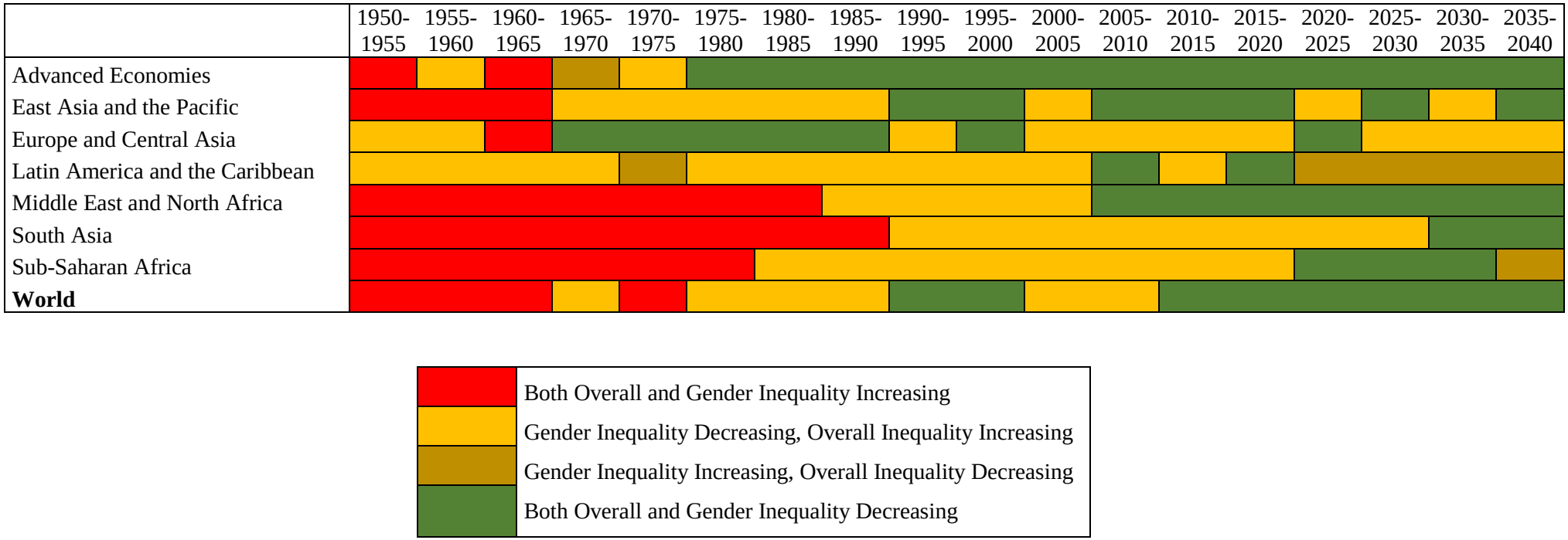

Figure 6. Summary of time periods during which trends in gender and overall education inequality go in simultaneous or different directions.

6. Summary and concluding remarks

The main aim of this research was to investigate the relationship between overall education inequality, its different subcomponents and the gender gap reversal in educational attainment. For that purpose, we have developed a new inequality index with nice decomposability properties adapted to the ordinal nature of educational attainment. Such decomposition – which for the first time allows articulating the different components of education inequality (inequality among women and men, inequality favoring men and inequality favoring women) into a coherent whole – can be extremely useful for scholars and policy-makers to gauge the transformations that certain factors might have exerted on different parts of the education distribution. Based on the Barro and Lee (2015) dataset we have calculated the values of the inequality index and the contribution of its components for 146 countries during a period spanning almost a century (1950-2040).

The composition of education inequality has changed dramatically during the period between 1950 and 2010. In the past, women used to be a much more homogeneous group compared to men in terms of education in several world regions. With education expansion, inequality among women quickly started to increase. We expected that as the gender gap in education reversed in favor of women, the inequality among them would be higher than the inequality among men. Surprisingly, our results show that inequality among men and women is now equally unequal in all world regions, and this is expected to remain so in the near future. Back in the 1950s, the educational advantage of men over women was by far the main contributor to education inequality in the world as a whole and in most of its regions, while the opposite was true about the educational advantage of women over men. Nowadays, men's educational advantage has become the least important contributor to education inequality in several regions, while women's educational advantage is the main contributor to education inequality in most high- and middle-income countries.

How have education inequality and the gender gap in education jointly evolved over time as education expands? Our findings suggest that most world regions have experienced, are experiencing or are expected to experience a period whereby both education inequality and the gender gap decline simultaneously – i.e. an opportunity window to potentially reap the benefits that such inequality reductions entail. Yet, assuming that large educational advantages of one sex over the other are normatively undesirable our findings suggest that in the near future there might be difficult trade-offs between gender and overall education inequality, as they both seem to run in opposite directions. If further education expansion contributes to further decreases in overall education inequality *and* increases in females' educational advantage over men (a scenario that appears likely until a ceiling of tertiary education attainment is reached in many countries), education planners might be facing a difficult ethical dilemma (particularly in high-income countries) upon which it will be necessary to reflect.

What should one expect for the more distant future? Will education inequality continue to decline and the gender gap to increasingly favor women? In line with Inglehart and Welzel (2005), our expectations are that as long as the modernization process continues to unfold, it will facilitate cultural changes that make gender equality increasingly likely. Since we

speculate that gender equality has come to stay, we expect that the gender gap in education will not continue increasing to attain the high levels it had in the initial stages of the male-dominated education expansion, but will rather hover around gender parity levels.

As regards education inequality there are a couple of reasons to *not* expect a continuation of its currently downward trend for a long time. First, while the global decline in overall education inequality (both for women and for men) is a cause of celebration pointing to an increasing equalization of opportunities across citizens worldwide, it is likely that new forms of education inequality not captured in our data might be (or already have been) replacing older ones. Indeed one can suspect that the observed declines in overall education inequality can be the attributed to the clustering of the educational attainment distribution at its top, which might be hiding an increasing diversity of superior education alternatives (like Masters and PhDs) not captured in the 7 categories of Barro and Lee's dataset. Some exploratory work carried out for the case of the US using census microdata samples (not shown here but available upon request) suggests that this might be indeed the case. When we enlarge the set of education categories from 7 to 9 (including Masters and PhDs), the decline in overall education inequality is less pronounced. In the same line, a recent study by Meschi and Scervini (2013) suggests that after a long period of sustained decline, education inequality is turning upwards in many European countries as a result of the expansion of tertiary and post-tertiary schooling.

Second, education inequality is intimately associated with technological progress and economic inequality, two forces that are unlikely to remain stable over time. Under the assumption that technology is skill-biased (Tinbergen 1974), technological progress will widen inequality among skill groups unless it is countered by increases in the supply of educated workers. Depending on the outcome of the 'race between education and technology' (Goldin and Katz 2008) – which is particularly uncertain – income and education inequality levels can vary to a great extent. Along similar lines, Milanovic (2016) recently suggested that the modern historical era from the preindustrial through the postindustrial period is characterized by the so-called 'Kuznets waves' of alternating increases and decreases in economic inequality, with many high-income countries currently in the upward portion of the wave. In all likelihood, the 'economic Kuznets waves' will translate into 'education Kuznets waves', thereby increasing education inequality as well. Whether or not such incipient trends will consolidate in the near future is a matter for future research.

References

Abul Naga, R. and Yalcin, T. (2008), "Inequality measurement for ordered response health data", *Journal of Health Economics* 27:1614-1625.

Allison, A. and Foster, J. (2004). Measuring health inequality using qualitative data. *Journal of Health Economics* 23: 505-524.

Ballarino, G., Bratti, M., Filippin, A., Fiorio, C., Leonardi, M. and Scervini, F. (2014). Increasing educational inequalities? Chapter 5 in Salverda, W., Nolan, B., Checchi, D.,

Marx, I., McKnight, A., Toth, I., van de Werfhorst, H. (Eds.) Changing inequalities in rich countries: Analytical and comparative perspectives. Oxford University Press.

Barro, R. and Lee, J. (2013), "A new data set of educational attainment in the world, 1950–2010", *Journal of Development Economics* 104:184-198.

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

Benaabdelaali, W., Hanchane, S., & Kamal, A. (2012). Chapter 13 Educational Inequality in the World, 1950–2010: Estimates from a New Dataset. *Inequality, Mobility and Segregation: Essays in Honor of Jacques Silber (Research on Economic Inequality, Volume 20)*, Emerald Group Publishing Limited, 337-366.

Castelló, A. and Domenech, R. (2002). Human capital inequality and economic growth: Some new evidence. *The Economic Journal*, 112: C180-C200.

Dorius, S. F., & Firebaugh, G. (2010).Trends in global gender inequality. *Social Forces*, *88*(5), 1941-1968.

Esteban, J. and Ray, D. (1994), "On the measurement of polarization", *Econometrica* 62:819-852.

Esteve, A., Schwartz, C., Van Bavel, J., Permanyer, I., Klesment, M. & Garcia, J. (2016). The end of hypergamy: Global trends and implications, forthcoming in *Population and Development Review*.

Farkas, G. (2003). Cognitive skills and noncognitive traits and behaviors in stratification processes. *Annual review of sociology*, 29(1), 541-562.

Härkönen, J. and Dronkers, J. (2006), "Stability and change in the educational gradient of divorce. A comparison of seventeen countries", *European Sociological Review* 22(5): 501-517.

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

Grant, M. and Behrman, J. (2010), "Gender gaps in educational attainment in less developed countries", *Population and Development Review* 36(1):71-89.

Inglehart, R. and Welzel, C (2005), *Modernization, cultural change and democracy. The human development sequence.* Cambridge University Press.

Jordá, V. and Alonso, J. (2016). New estimates of educational attainment using a continuous approach (1970-2010). *World Development* (DOI: 10.1016/j.worlddev.2016.10.005).

Klasen, S. and Lamanna, F. (2009). The impact of gender inequality in education and employment on Economic Growth: New evidence for a panel of countries. *Feminist Economics*, 15(3):91-132.

Kobus, M. and Milos, P. (2012), "Inequality decompositions by population subgroups for ordinal data", *Journal of Health Economics* 31:15-21.

Lafortune, J. (2013). Making yourself attractive: Pre-marital investments and the returns to education in the marriage market. *American Economic Journal: Applied Economics*, 5(2), 151-178.

Lee, R. and Mason, A. (2006) "What is the Demographic Dividend?", *Finance and Development* (September) 16-17.

Lutz, W. & KC, S. (2011). Global human capital: Integrating education and population. *Science*, 333:587-592.

Meara, E. R., Richards, S., & Cutler, D. M. (2008). The gap gets bigger: changes in mortality and life expectancy, by education, 1981–2000. *Health Affairs*, 27(2), 350-360.

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

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

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

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

Psacharopoulos, G., & Patrinos, H. A. (2004). Returns to investment in education: a further update. *Education economics*, 12(2), 111-134.

Sahn, D. and Younger, S. (2007). Decomposing world education inequality. Cornell Food and Nutrition Policy Program No. 187. Ithaca, NY: Cornell University.

Tinbergen, J. (1974), "Substitution of graduate by other labor", *Kyklos* 27(2):217-226.

Van Bavel, J. (2012), "The reversal of gender inequality in education, union formation, and fertility in Europe", *Vienna Yearbook of Population Research* 10:127-154.

Van Bavel, J. and Klesment, M. (2017), "Educational Pairings, Motherhood, and Women's Relative Earnings in Europe", *Demography*

Wils, A. and Goujon, A. (1998), "Diffusion of education in six world regions, 1960-90", *Population and Development Review* 24(2):357-368.

Appendix 1. Derivation of Equation [3].

Manipulating algebraically it is easy to show that

$$I(p_1,\cdots,p_k)=\sum_{i=1}^{i=k}\sum_{j=1}^{j=k}p_ip_j\mathbb{I}(i,j)=\frac{\sum_{i=1}^{k}\sum_{j\neq i}N_iN_j}{N(N-1)}=\frac{\sum_{i=1}^{k}\sum_{j\neq i}(N_i^f+N_i^m)(N_j^f+N_j^m)}{N(N-1)}$$

$$=\frac{\sum_{i=1}^{k}\sum_{j\neq i}N_i^fN_j^f+\sum_{i=1}^{k}\sum_{j\neq i}N_i^mN_j^m+\sum_{i=1}^{k}\sum_{j\neq i}N_i^fN_j^m+\sum_{i=1}^{k}\sum_{j\neq i}N_i^mN_j^f}{N(N-1)}$$

$$=\left[\frac{N^f(N^f-1)}{N(N-1)}\right]\frac{\sum_{i=1}^{k}\sum_{j\neq i}N_i^fN_j^f}{N^f(N^f-1)}+\left[\frac{N^m(N^m-1)}{N(N-1)}\right]\frac{\sum_{i=1}^{k}\sum_{j\neq i}N_i^mN_j^m}{N^m(N^m-1)}+$$

$$+\left[2\frac{N^fN^m}{N(N-1)}\right]\frac{\sum_{i=1}^{k}\sum_{j\neq i}N_i^fN_j^m}{N^fN^m}$$

$$=s_fI_f+s_mI_m+\frac{s_b}{N^fN^m}\left(\sum_{i=2}^{i=k}\sum_{j<i}N_i^fN_j^m+\sum_{i=2}^{i=k}\sum_{j<i}N_i^mN_j^f\right)$$

$$=s_fI_f+s_mI_m+s_b\left(\sum_{i=2}^{i=k}\sum_{j<i}p_i^fp_j^m+\sum_{i=2}^{i=k}\sum_{j<i}p_i^mp_j^f\right)$$

$$=s_fI_f+s_mI_m+s_b\left(A_f+A_m\right).$$

This is the decomposition we were looking for.

Appendix 2. Actual and predicted attendance of educational stages for men 1950-2040