What Explains the Gender Gap Reversal in Education?

Theory and Evidence

Laurent Bossavie
Ohto Kanninen
What Explains the Gender Gap Reversal in Educational Attainment?

Theory and Evidence *

Laurent Bossavie† and Ohto Kanninen‡

Keywords: Educational attainment, gender gap, gender differences in test score distribution.

JEL-Codes: I20, J16

*We are grateful to Andrea Ichino and Luigi Guiso for constant advice and discussions. We also would like to thank Jerôme Adda, Christian Dustmann, Nicole Fortin, Paola Giuliano, Peter Hansen, Steve Machin, Tommaso Naniccin, Tuomas Pekkarinen, Kjell Salvanes and Olmo Silva as well as all participants to the micro-econometrics working group at the EUI for helpful comments and suggestions. The views expressed in the paper are those of the authors and do not necessarily represent those of the World Bank or of the Governments they represent.

†The World Bank. Corresponding author. E-mail: lbossavie@worldbank.org. Address: The World Bank, 1818 H Street NW, 20433 Washington DC, USA. Phone: 202-751-6478.

‡Labour Institute for Economic Research (LIER), Helsinki.
Abstract

The reversal of the gender gap in educational attainment is becoming a global phenomenon. Its drivers, however, are not well understood and remain largely untested empirically. This paper develops a unified conceptual framework that allows to formulate and test two main hypotheses for the reversal. It introduces the tail hypothesis, which builds on the lower dispersion of academic achievement among females observed globally. It also formulates the mean hypothesis, which states that females' average school performance and returns to education have increased over time relative to males'. The paper theoretically shows that both hypotheses can explain the reversal in our framework. The parameters of the two hypotheses derived from the model are then estimated using educational attainment data from 1950 to 2010 in more than 100 countries. The paper finds that both hypotheses strongly predict the gender gap dynamics in educational attainment when estimated separately. When jointly estimated, however, the predictive power of the tail hypothesis is more robust, while the mean hypothesis plays a weak role in secondary school completion dynamics.

This paper is a product of the Social Protection and Labor Global Practice Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The authors may be contacted at lbossavie@worldbank.org.
1 Introduction

The dramatic increase in educational attainment over the past decades was accompanied by a striking and puzzling phenomenon. As individuals stayed longer in school over time, females not only caught up with males’ education levels, but progressively attained higher levels of schooling. This phenomenon, sometimes referred to as the gender gap reversal in education, is observed in virtually all high-income countries and in a growing majority of low and middle-income countries. Although the reversal is becoming close to universal, its origins are not well understood. While a lot of attention has been devoted to explaining the gender gap in labor market outcomes, the gender gap reversal in educational attainment has been understudied. A few contributions, confined to the US context, have proposed potential explanations for this phenomenon (Goldin et al. 2006, Cho 2007, Chiappori et al. 2009 and Fortin et al. 2015). However, to the best of our knowledge, there has been no attempt to model and test alternative hypotheses for this reversal in a unified theoretical framework. This paper aims at filling this gap.

Understanding the origins of the educational gender gap reversal is important in itself, but also to help explaining gender dynamics in other areas, particularly in the labor market. For policy purposes, it is also important to identify whether differences in educational outcomes between genders originate from distortions such as social barriers and discrimination or, instead, from optimizing behaviors based on potential gender differences in preferences or traits. Learning about the origins of the gender gap reversal can help determine whether policy interventions are needed to address the growing disadvantage of boys in school, and identify potential areas of interventions.

The paper starts by establishing that the gender gap reversal in education is becoming a global phenomenon. Using time-series on enrollment and completion rates in education by gender in more than 100 countries, we show that the reversal occurred in virtually all high-income countries, as well as in a rapidly growing proportion of lower-income countries. We also evidence that the reversal is broader than the reversal in tertiary enrollment evidenced by previous literature, and also occurred in primary and secondary school completion. While females have historically rep-

\[1\] For a summary of evidence on gender differences in preferences, see Croson and Gneezy (2009).
\[2\] The latter applies mostly to lower-income countries.
resented the minority of tertiary, secondary and primary education completers, they have become the majority over time.

The first contribution of the paper is to develop a simple conceptual framework to account for the gender gap reversal in education. In this setting, differences in the distributions of academic achievement between genders play a central role. The model builds on the micro foundations of investment in schooling of Card (1994), before aggregating individual schooling decisions to the macro level. It has three building blocks. First, at the micro level, the optimal number of years of schooling attained by individuals is increasing in their level of academic achievement. Second, the benefits of investing in schooling are allowed to vary over time and between genders. Third, we introduce two alternative assumptions regarding the distribution of academic achievement by gender. These distributional assumptions, together with an increase in the economy-wide benefits of investing in schooling, generate either a tail dynamics or a mean dynamics, depending on the assumption made. Our model allows to combine these two dynamics within a unified theoretical framework.

Another contribution of the paper is to propose and model the tail dynamics hypothesis, or tail hypothesis, and to show that it can generate the gender gap reversal in education. A key result of our model under the tail hypothesis is that a higher variance in academic achievement among males is sufficient to produce the gender gap reversal, independently of the relative mean achievement of males and females. We show that the tail hypothesis implies first a gender gap reversal in favor of females, but ultimately gender equality in educational attainment as the net benefits of education increase over time and enrollment becomes universal. The hypothesis is novel in the literature and builds on the findings of Machin and Pekkarinen (2008) who find that girls exhibit a lower variance in mathematics and reading test scores relative to boys in virtually all OECD countries. It also formalizes the intuition of Becker et al. (2010b), who suggest that the lower variance of non-cognitive skills among females can generate a higher elasticity of enrollment to returns to schooling.

---

3In the remainder of the paper, we use academic achievement to refer to scholastic performance measured by scores obtained in standardized tests, while educational attainment refers to a given level of schooling attained by an individual, e.g. secondary completion.

4There also exists an abundant literature in psychology reporting similar findings for different dimensions of abilities or traits, such as Frasier (1919), Hedges and Nowell (1995), or Jacob (2002), among many others.
We also model the mean dynamics hypothesis within our framework. Contrary to the tail hypothesis, the predictions of the mean hypothesis depend directly on the development of gender-specific means, and gender equality in educational attainment is only achieved if gender-specific means are the same. The mean dynamics hypothesis can be modeled in two alternative ways within our framework. It first can be formalized through the mean benefit hypothesis (MBH), which posits that the net benefits of schooling for females have increased more rapidly over time than for males (Goldin et al. 2006, Chiappori et al. 2009). Second, it can be modeled through the mean performance hypothesis (MPH), which claims that the mean performance of females in achievement tests has increased over time relative to males’ (Cho 2007). We show that both mean hypotheses are algebraically similar in our framework, and can both produce a reversal of the gender gap in educational attainment.

Finally, another contribution of the paper is to empirically estimate the contributions of the two hypotheses to the gender gap reversal. distinguish between the two hypotheses and to quantity their respective contributions to the gender gap reversal. This is, to the best of our knowledge, the first paper that formally quantifies the contributions of alternative hypotheses within a unified framework. Using our model, we show that each of the mean and tail dynamics hypothesis can be expressed by a reduced-form equation and summarized by two key parameters. We then fit each hypothesis using harmonized time-series data on educational attainment by gender from 1950 to 2010 in a large sample of countries. We are able to estimate each hypothesis’ parameters derived from our theoretical framework first separately, and then jointly.

We find that each of the mean and tail dynamics hypothesis parameters strongly predicts the dynamics of the gender gap over time in tertiary enrollment. However, accounting for both parameters in the estimation strongly increases predictive power, as opposed to specifications with only one of the two hypotheses parameters. For secondary school completion, however, the predictive power of the mean hypothesis parameter drastically declines once the tail hypothesis parameter also enters the estimation. In contrast, the magnitude of the tail hypothesis parameter is quite robust to the inclusion of the mean hypothesis in both the tertiary enrollment and secondary school completion estimations.
Our findings suggest that while the tail and mean dynamics have both been at play in the gender gap reversal in education, the tail hypothesis has a strong and independent explanatory power that has not been uncovered by previous literature. Accounting for the tail hypothesis significantly increases the predictive power for tertiary enrollment dynamics, and it has stronger and more robust explanatory power for secondary school completion compared to the mean hypothesis. Our model estimates for the tail hypothesis parameter are also quite close to the male-to-female variance ratio in test scores estimated from international student assessments. This provides further comfort on the validity of the tail dynamics hypothesis.

The paper is organized as follows. Section 6.4 presents empirical facts on the gender gap reversal in education motivating the paper. Section 3 lays out our conceptual framework. Section 4 formulates two main hypotheses for the gender gap reversal within this framework. Section 5 describes our empirical methodology to estimate the contribution of the two hypotheses to the gender gap reversal. Section 6 presents our main empirical results and discusses their implications. Section 7 concludes.

2 Empirical Motivation

2.1 Data

The data used throughout the paper is from the Barro-Lee educational attainment database. The dataset consists of harmonized country-level data on educational attainment constructed from nationally representative surveys and census data. The data is reported at five-year intervals from 1950 and 2010. It is also disaggregated into five-year age groups. The availability of age-specific data is valuable in our context, as it allows to construct the dynamics of educational attainment of specific age cohorts over time. Educational attainment is also reported separately for males and females, which allows to follow the evolution of educational attainment across cohorts by gender. Time-series of educational attainment are available for a total of 146 countries in the database,
which include 36 OECD economies and 110 non-OECD countries.\textsuperscript{5} The dataset reports information on the distribution of educational attainment of the adult population at seven different levels of schooling: no formal education, incomplete primary, complete primary, lower secondary, upper secondary, incomplete tertiary, and complete tertiary.\textsuperscript{6}

To describe time trends by gender, we focus on three main levels of educational attainment: primary school completion, secondary school completion, and enrollment in tertiary education. As the data is reported every five years from 1950 to 2010, we observe a maximum of 13 time variant observations per country, for a given level of educational attainment. We exclude eight countries from the original Barro-Lee dataset, due to highly inconsistent data and missing values on educational attainment.\textsuperscript{7}

We compute completion and enrollment rates for 5-year-band birth cohorts born from 1921 to 1990, by gender.\textsuperscript{8} We measure the enrollment rate in tertiary education at time $t$ as the share of the population age 25-29 that has ever enrolled in tertiary education. The secondary education completion rate is measured as the share of the population age 20-24 at time $t$ that has completed secondary education and above. Finally, primary education completion rates are measured as the share of the population age 15-19 at time $t$ that has completed primary education and above.\textsuperscript{9} The female-to-male ratio is measured as the enrollment or completion rate of females over that of males for a given level of educational attainment.

Table I displays the summary statistics of enrollment and completion rates for primary, secondary and tertiary education in OECD and non-OECD countries from the Barro-Lee database in 1950 and 2010. As documented in Barro and Lee (2013) and Barro and Lee (2016), the data shows a large increase in educational attainment over the period covered by the data. Meanwhile, and more importantly in the context of the paper, it also shows that the enrollment and completion rates have

\textsuperscript{5}In case of missing observations over time, data points were obtained by forward and backward extrapolation of the census/survey observations on attainment. See Barro and Lee (2013) and Barro and Lee (2016) for a complete description of the dataset construction and methodology.

\textsuperscript{6}Results for the unbalanced panel of countries are also reported in the Appendix.

\textsuperscript{7}Excluded countries are Botswana, Burundi, Cameroon, Gabon, Guyana, Mozambique, Libya and Yemen.

\textsuperscript{8}For primary education, we compute completion rates for individuals born between 1936 to 1995, from 1931 to 1990 for secondary school completion and from 1921 to 1980 for tertiary education.

\textsuperscript{9}The choice of these age brackets was motivated by the fact that individuals would be expected to have enrolled in or completed the corresponding level of education by these ages, if they ever did.
increased more rapidly for females than for males for all three levels of educational attainment. This pattern is observed for both OECD countries and non-OECD countries.

2.2 The Gender Gap Reversal in Educational Attainment

Several papers have reported a convergence followed by a reversal in the relative number of females attending tertiary education in the US (Charles and Luoh 2003, Goldin et al. 2006, Chiappori et al. 2009, Becker et al. 2010a, Becker et al. 2010b, Autor and Wasserman 2013). In table II and Figure I, we look at the evolution of the female-to-male ratio in tertiary enrollment worldwide, by grouping country-level data into seven world regions. As shown in Column 5 and 6 of Table II, young males in 1950 used to outnumber females among participants to tertiary education in virtually all countries in the sample (92%). In contrast, by 2010, 73% of countries in the sample had seen their gender imbalances in tertiary enrollment revert from a male majority to a female majority. Even among non-OECD countries, 66% had experienced the gender gap reversal in tertiary enrollment by 2010. The gender gap reversal in tertiary enrollment is therefore clearly becoming a global phenomenon. In advanced economies, Europe and Central Asia and Latin America, the share of countries where females outnumber males among tertiary education students is over 90 percent. In contrast, most countries in South Asia and Sub-saharan Africa have yet to experience the gender gap reversal in tertiary enrollment.

When averaged over all countries in the sample, the Barro-Lee data show a reversal in the gender gap in tertiary enrollment, as reported in Panel A of Figure I. For the earlier cohorts born in the 1920s, the average female-to-male ratio in tertiary enrollment for all countries in the sample was around 0.40, meaning that for ten females enrolled in tertiary education, only four females was enrolled. This ratio was fairly similar across all regions of the world. In contrast, among the most recent cohorts, the average female-to-male ratio had reached 1.22 globally, meaning that there were about 22% more females enrolled in tertiary education than males worldwide.

10 Pekkarinen (2012) has also reported this phenomenon for Scandinavian countries.
11 This regional grouping follows the World Bank classification of world regions and includes: advanced economies, Latin American and the Caribbean, Europe and Central Asia, East-Asia an Pacific, the Middle-East and North Africa, South Asia and Sub-saharan Africa.
Out of the seven world regions, five have already crossed the horizontal line of a female-to-male ratio of 1, meaning that they experienced the gender gap reversal in tertiary enrollment. Sub-Saharan African and South Asia are the only two regions that have yet to experience the reversal, with a female-to-male ratio in tertiary enrollment remaining below one, despite a noticeable increase across cohorts.

Panel A of Figure I also shows that the timing of the reversal varies across regions. The gender gap reversal in tertiary enrollment occurred first in Europe and Central Asia for cohorts born at the beginning of the 1950s. In advanced economies and Latin America American countries, this reversal occurred first for cohorts born around 1960. Among countries in East Asia and the Middle-East and Northern Africa, the reversal only took place among cohorts born in the late 1970s.

Table II and Figure I show that the global gender gap reversal in educational attainment is not limited to tertiary education enrollment. The gender composition of secondary school completers has also reversed from a male majority to a female majority over time. While young females used to outnumber males among secondary school non-completers in only 6% of the countries in the sample in 1950, females now represent the majority of secondary school completers in 71% of countries in the sample. As for tertiary enrollment, the gender gap reversal in secondary school completion already occurred in five out of the seven world regions. The share of countries that have experienced the reversal is highest in advanced economies (92%) and lowest in Subsaharan Africa (38%). As for tertiary enrollment, Europe and Central Asia is the first region to have experienced the gender gap reversal in secondary school completion.

Finally, we also observe a closure and reversal of the gender gap in primary school completion, although patterns are less clear cut than for secondary and tertiary education. As shown in Panel C of Figure I, the female-to-male ratio in primary completion averaged over all countries in the sample was 0.82 for the oldest cohorts born at the beginning of the 1930s, but had reached a mean of 1.03 among cohorts born at the beginning of the 1990s. When disaggregated by regions, all regions have reached female-to-male ratios in primary school completion around or slightly above one for the latest cohorts, with the exception of South Asia. Similarly, a majority of countries have experienced the gender gap reversal in primary school completion in most regions (Table II). The
magnitude of the reversal is however smaller than for secondary school completion and tertiary enrollment. Among advanced economies in particular, the share of countries that experienced the gender gap reversal in primary school completion has hardly changed from 1950 to 2010.

The fact that the gender gap reversal is not as clear cut for primary school completion as it is for secondary and tertiary education can be explained by several factors. First, in many countries in our sample, primary school completion has been close to universal already for some time and, as a result, there has been little variation in the gender ratio which remained around one over our period of study. This is the case for advanced economies, Europe and Central Asia and Latin America (Figure I). A related explanation is that primary schooling has been mandatory for some time for many countries in the sample, which leaves little room for variation in the gender ratio in primary school completion over time. As result, primary school completion is mainly driven by institutional features rather than optimal economic decisions of investment in schooling by gender.

3 Theoretical Framework

3.1 Micro Foundations

We now develop a conceptual framework that allows to formulate alternative hypotheses for the gender gap reversal. The economy is assumed to be populated by of mass individuals that differ in their level of academic achievement $z$, which is continuous and perfectly observed by individuals. For the sake of simplicity, a single-period model is assumed where individuals receive the benefits of their investment in schooling in the same period as they invest. Individuals choose years of schooling $s$ to maximize their level of utility $U$. Building on Card (1994), we simply express the utility function of individuals in the economy as:

$$U(s) = B(s) - C(s),$$

12 Academic achievement can be interpreted as the product of a complex combination of individual ability as well as other inputs such as effort and motivation, which can empirically be measured by test scores (Heckman and Kautz, 2012)
where $B(s)$ is the benefit function of schooling, with $B'(s) > 0$ and $B''(s) < 0$. $C(s)$ is the cost function of schooling, which is assumed to be increasing and convex in $s$. The first-order conditions for the individual maximization problem read as:

$$B'(s) = C'(s),$$

where $B'(s)$ and $C'(s)$ denote the marginal benefits and costs of schooling, respectively. Following Card (1994), we linearize the model by assuming that $B'(s)$ and $C'(s)$ are linear functions of $s$, with $B'(s)$ having an individual-specific intercept:

$$B'(s) = z_j - k_1 s,$$

$$C'(s) = k_2 s,$$

where $k_1 > 0$ and $k_2 > 0$. Intuitively, individuals with higher levels of academic achievement $z_j$ get higher marginal benefits from schooling.\(^{13}\) In this framework, the optimal level of schooling $s$ chosen by individual $j$ can be expressed as:

$$s_{j}^* = z_j \cdot b,$$  \hspace{1cm} (1)

where $b = \frac{1}{k_1 + k_2}$ is an exogenous technology parameter, which we call net benefits of schooling (hereafter: net benefits). It is identical for all individuals in the economy. $b$ captures the monetary benefits, non-monetary benefits and costs of schooling in the economy. The optimal level of schooling chosen by individual $j$ is therefore strictly increasing in individual academic achievement $z_j$.

In this framework, the minimum level of academic achievement $\bar{z}$ so that individuals attain a given level of schooling $\bar{s}$, such as tertiary education, can be expressed as:

$$\bar{z} = \frac{\bar{s}}{b}$$  \hspace{1cm} (2)

\(^{13}\)In the original model of Card (1994), the expression for $C'(s)$ also includes an individual-specific intercept, capturing individual-specific circumstances such as access to wealth and network or taste for education. For the sake of simplicity, we abstract from this distinction in our model.
Equation 2 states that individuals whose academic achievement is below the threshold $\bar{z}$ attain an optimal level of schooling below $\bar{s}$, while individuals whose achievement is equal to or greater than $\bar{z}$ attain $\bar{s}$ and above. It also implies that the achievement threshold is determined by $b$, common to all individuals in a given cohort. An immediate implication of Equation 2 is:

$$\frac{\partial \bar{z}}{\partial b} \leq 0.$$ 

In words, the minimum level of academic required to attain a given level of schooling $\bar{s}$, such as tertiary education, decreases with the net benefits of investing in schooling in the economy.

### 3.2 Aggregate Educational Attainment and Gender Ratio

We now assume that the economy is populated by successive cohorts. Each cohort comprises a continuum of agents that invest in schooling and differ in their level of academic achievement $z$. Let $f_z(\bar{z})$ denote the probability density function of academic achievement $z$ in the population of a given cohort. The complementary cumulative distribution function of $z$ can be expressed as:

$$G_z(\bar{z}) = \int_\bar{z}^{+\infty} f_z(\bar{z}) \, d(\bar{z}).$$

All individuals belonging to the same cohort are exposed to the same value of the exogenous parameter $b_t \equiv \frac{1}{k_{1,t} + k_{2,t}}$, irrespective of their academic achievement. Given the micro properties of the model summarized in Equation (1) and (2), the share of individuals in the cohort that attain a level of schooling of at least $\bar{s}$ is given by:

$$P(\bar{z}) = 1 - F_z(\frac{\bar{s}}{b}) = G_z(\frac{\bar{s}}{b}) = G_z(\bar{z}). \quad (3)$$

Equation (3) states that the mass of individuals that attain a level of schooling of at least $\bar{s}$ or higher, for a given value of $b$, is made of all individuals whose academic achievement is above the achievement threshold $\bar{z}$. Exogenously to individual schooling decisions, $b$ varies across cohorts.
Equation (3) implies: \[
\frac{\partial P(\bar{z})}{\partial b} \geq 0.
\]
The share of individuals that attain a level of schooling equal or higher than \( \bar{s} \) therefore increases with the net benefits of schooling \( b \). One key implication at the aggregate level is that, as the net benefits of schooling rise, educational attainment increases and the mean academic achievement of individuals that attain a given level of schooling decreases. In Appendix A1.3, we show that this result holds empirically, using US data.

We now introduce the possibility that \( G_z(\cdot) \) and \( \bar{z} \) differ between genders. Let \( z_m \) and \( z_f \) be random variables that denote the academic achievement of males and females, respectively. The academic achievement threshold \( \bar{z} \) is also allowed to vary by gender, and is denoted \( \bar{z}_m \) and \( \bar{z}_f \) for males and females, respectively. In our framework, the fraction of the population in the cohort that attains a given level of schooling \( \bar{s} \), assuming the total population of males and females is the same, is given by:

\[
P(\cdot) \equiv \frac{G_{z_f}(\bar{z}_f) + G_{z_m}(\bar{z}_m)}{2}
\]

and the female-to-male ratio among individuals that choose a given level of schooling, denoted \( R(\bar{z}) \), can be expressed as:

\[
R(\cdot) \equiv \frac{G_{z_f}(\bar{z}_f)}{G_{z_m}(\bar{z}_m)}
\]

The implications of our model at the aggregate level are fully summarized by Equation 4 and 5. In the remainder of the paper, \( G_{z_f} \) and \( G_{z_m} \) are assumed to be characterized by their first two moments, where \( \mu \) and \( \sigma^2 \) denote the mean and variance of the distributions, respectively.

4 Two Main Hypotheses for the Gender Gap Reversal

Within this theoretical framework, we formulate two main hypotheses for the gender gap reversal: the tail dynamics hypothesis (or tail hypothesis), and the mean dynamics hypothesis (or mean hypothesis). The mean dynamics hypothesis itself can be formalized in two alternative ways:
through the mean performance hypothesis (MPH), or through the mean benefit hypothesis (MBH).
As shown in Section A3.2, these two sub-hypotheses are algebraically similar within our theoretical framework. We therefore group them together under the mean dynamics hypothesis. Table V provides a summary of the underlying assumptions of these different hypotheses.

The tail hypothesis states that $G_z(\cdot)$ is gender specific with $\sigma_m > \sigma_f$. In words, the academic achievement of males and females are characterized by different distributions, and in particular a greater variance among males. The net benefits of schooling $b$ are assumed to be identical between genders. The increase in $b$ over time (or, equivalently, the decrease in the academic achievement threshold $\bar{z}$) combined with a greater dispersion of academic achievements among males, produces the reversal in educational attainment.

The mean benefit hypothesis (MBH) claims that the net benefits of schooling differ between genders, i.e. there exist gender-specific $\bar{z}_f$ and $\bar{z}_m$ (or, equivalently in our framework, $b_f$ and $b_m$) that have different dynamics over time. Prior to the reversal, $\bar{z}_f > \bar{z}_m$ (or, equivalently in our framework, $b_f < b_m$) before $\bar{z}_m$ progressively converges towards $\bar{z}_f$ and surpasses it over time, generating the reversal. The distributions of academic achievement $G_z(\cdot)$ are assumed to be identical for males and females.

The mean performance hypothesis (MPH) claims that the average academic achievement of females $\mu_f$ has increased more rapidly than for males over time and ultimately surpassed $\mu_m$, producing the gender gap reversal. The variance of $z$ and the net benefits of schooling (or, equivalently, the academic achievement threshold $\bar{z}$) are assumed to be identical for both genders.

Among these two mean dynamics hypotheses, we henceforth focus on the mean performance hypothesis, for reasons discussed in section 4.2.
4.1 The Model under the Tail Dynamics Hypothesis

The overwhelming majority of the literature on gender differences in scholastic achievement has focused on the gap in mean performance between boys and girls, especially in mathematics.\textsuperscript{14} Machin and Pekkarinen (2008), however, also find systematic differences in the variance of test scores between genders. They show that the variance of boys is greater than that of girls in virtually all OECD countries.\textsuperscript{15}

In Figure II, we show that this findings holds when using the most recent available wave of the Project for Student International Assessment (PISA) data from 2015, which consists of a larger sample of countries. PISA tests nationally representative samples of 15-year-old students in mathematics and reading, and test scores are comparable across countries. In the our sample of 67 countries, the average ratio of the male test score variance over the female test score variance is 1.17 for both reading and mathematics test scores.\textsuperscript{16} The variance ratio ranges from 0.96 in Algeria to 1.62 in Jordan for reading, and from 0.90 in Algeria to about 1.50 in Argentina and Jordan. The test score variance is higher than that of females’ in 65 countries for mathematics and in 64 countries for reading, and the gender difference is statistically significant in 58 countries for both reading and mathematics.\textsuperscript{17} In line with this evidence, Becker et al. (2010a) and Becker et al. (2010b) suggested that a lower variance of non-cognitive skills among females could explain the gender gap reversal in education by inducing a higher elasticity of females’ schooling to returns to education.

Understanding the higher dispersion of boys’ test scores is beyond the scope of this paper, but it could be linked to another growing strand of literature suggesting that boys’ achievement is more sensitive to inputs in early childhood and at school. Bertrand and Pan (2013) document that boys

\textsuperscript{14}See for example Guiso et al. (2008), Niederle and Vesterlund (2010), Dee (2007), Fryer and Lewitt (2010) or Dossi et al. (2019) among many others.

\textsuperscript{15}Test scores can here be thought as the empirical measure of individual academic achievement $z$ in our theoretical model, which is the observable product of a complex combination of cognitive and non-cognitive skills, as well as other inputs such as effort and motivation (Heckman and Kautz, 2012).

\textsuperscript{16}In the smaller sample of 40 OCED countries from PISA 2003 used by Machin and Pekkarinen (2008), the average male-to-female variance ratio is 1.20 for reading and 1.21 for mathematics.

\textsuperscript{17}Although this evidence is recently known to economists, long-standing evidence on the larger variance in the distribution of some traits among males has been reported in the psychology literature, such as Frasier (1919), Hedges and Nowell (1995) or Jacob (2002), among many others.

14
raised in single-parent families have twice the rates of behavioral and disciplinary issues as boys raised in two-parent families; Autor et al. (2015) show that the boy-girl gap in kindergarten readiness, test scores, truancy, disciplinary problems, disability, juvenile delinquency, and high school graduations is larger in more disadvantaged families. Fan et al. (2015) report that maternal employment during early childhood reduces boys’ eventual educational attainment relative to that of girls. With respect to the impact of school inputs, Autor et al. (2016) find that boys benefit more than girls from cumulative exposure to higher quality schools. Legewie and DiPrete (2012) also show that boys’ development or inhibition of anti-school attitudes and behavior is much more sensitive to peer socio-economic status, which is correlated with school quality measures. Finally, Cornelissen and Dustmann (2019) show that receiving additional schooling before age five has stronger effects on the cognitive and non-cognitive outcomes of boys from disadvantaged backgrounds relative to girls.

We then formulate the tail hypothesis within our theoretical framework. Let \( z_m \) and \( z_f \) be random variables that denote the academic achievement of males and females, respectively, and let \( f_z(\bar{z}_m) \) and \( f_z(\bar{z}_f) \) denote their density functions. The tail hypothesis assumes that \( \sigma_m > \sigma_f \), where the distribution of \( z_m \) and \( z_f \) are assumed to be invariant over time. Importantly, no assumption is imposed on the relative value of \( \mu_m \) and \( \mu_f \). For empirical estimations and illustrations, we assume that \( z_m \) and \( z_f \) are normally distributed, as it is the closest approximation of empirical test score distributions.\(^{18}\)

In our framework, the share of individuals choosing a level of schooling of at least \( \bar{s} \) among the population of a given cohort is given by:

\[
P^{TH}(\cdot) = \frac{G_{z_f}(\bar{z}, \mu_f, \sigma^2_f) + G_{z_m}(\bar{z}, \mu_m, \sigma^2_m)}{2},
\]

and the female-to-male ratio among individuals that attain a level of schooling of at least \( \bar{s} \), denoted \( R(\bar{z}) \), can be expressed as:

\[
R^{TH}(\cdot) = \frac{G_{z_f}(\bar{z}, \mu_f, \sigma^2_f)}{G_{z_m}(\bar{z}, \mu_m, \sigma^2_m)}
\]

\(^{18}\)However, the key results of the model under the tail hypothesis also hold under alternative distributions. Formal proofs for alternative distributions are available upon request.
Panel A of Figure IV displays the distributions of academic achievement for both genders when \( \sigma_m^2 > \sigma_f^2 \), together with the achievement threshold \( \bar{z} \) that truncates the distributions of those that attain a level of schooling of at least \( \bar{s} \). Individuals whose academic achievement is to the right side of the achievement threshold attain at least \( \bar{s} \), while those to the left of the achievement threshold do not. Panel B displays the two corresponding complementary cumulative distribution functions (CCDFs) under this assumption, where the reversal occurs when the two CCDFs cross. Figure V illustrates the relationship between the share of individuals choosing a level of schooling of at least \( \bar{s} \) - referred to as the enrollment or completion rate hereafter - and the gender ratio among those choosing at least \( \bar{s} \). The reversal occurs when the ratio \( \frac{G_{zf}(\bar{z})}{G_{zm}(\bar{z})} \) reaches values above 1.

For the reversal to occur under the tail hypothesis, a decrease in the academic achievement threshold \( \bar{z} \) or, equivalently, an increase in the net benefits of education \( b \) which translates into an increase in \( P \), is required. Our distributional framework of investment in schooling also implies that the mean achievement of individuals that attain a given level of schooling \( \bar{s} \) decreases as the share of population that attains \( \bar{s} \) increases. In Appendix A1.2, we show that the average academic achievement of individuals attending tertiary education has indeed decreased over time, using US data.

Under the assumption that \( \sigma_m^2 > \sigma_f^2 \), it can be shown that the relationship between the female-to-male ratio among individuals choosing a level of schooling of at least \( \bar{s} \), denoted \( R^{TH} \), and the share of individuals that achieve a level of schooling of at least \( \bar{s} \), denoted \( P^{TH} \), exhibit three key properties:

**Proposition 1.** \( R^{TH} \) tends to zero when \( P^{TH} \) tends to zero.

**Proposition 2.** \( R^{TH} \) tends to one when \( P^{TH} \) tends to one.

**Proposition 3.** There exists a value of \( P^{TH} \in [0, 1] \) such that \( R^{TH} = 1 \). This value is unique and always exists.

**Proof.** See Appendix A3.

The distributional assumption \( \sigma_m^2 > \sigma_f^2 \) is a necessary and sufficient condition for Proposition 1
to 3 to hold, without any condition imposed on the relative values of $\mu_m$ and $\mu_f$. In addition, Proposition 1 to 3 also hold when $z$ is assumed to follow two-parameter probability distribution functions other than the normal distribution.19

4.2 The Model under the Mean Dynamics Hypothesis

It can be shown that the mean benefit and mean performance hypotheses are algebraically similar in our framework, under certain conditions (Appendix A3.2). Likewise, Goldin et al. (2006) argue that the relative increase in females’ performance in upper secondary school in the US was linked to a disproportionate increase in the returns to college education for women, suggesting that both phenomenon are closely tied. In the remainder of the paper, we focus on the mean performance hypothesis as it is more straightforward to estimate empirically. The mean benefit hypothesis is however presented in more detail in Appendix A2.1.

A change over time in the distribution of $z$ for females relative to males, in particular a relative increase in females’ mean academic achievement, can generate a reversal in the gender gap in our framework. Such changes over time have been highlighted by Goldin et al. (2006), Cho (2007), Fortin et al. (2015) or Yamaguchi (2018) for the US, and suggested as a potential explanation for the reversal. Goldin et al. (2006) use samples of high school graduating seniors from three US longitudinal surveys in 1957, 1972, 1988, among which the last two are nationally representative. Looking at achievement test scores, they find that girls reduced their disadvantage in math, and increased their advantage in reading between 1972 and 1992.

Using three nationally representative longitudinal data sets of high school students, Cho (2007) also looks at the evolution of female performance in high school over time. He reports that females’ mean test scores in high school have increased more rapidly than males’ over the past three decades. Using a simple Oaxaca decomposition, he finds that women’s progress in high school achievement can account for more than half of the change in the college enrollment gender gap over the past thirty years.

19Proofs for distribution functions other than the normal distribution are available upon request.
Fortin et al. (2015) use self-reported grades, rather than test scores, to look at the evolution of female high-school performance over time. Using a sample of 12th graders from the Monitoring the Future (MTF) study in 1976 and 1991, they report that the gender gap in the mean grades of high school seniors remained very stable since the 1970s. They find, however, that the mode of girls has shifted upwards over the period 1980–2010 compared to that of boys’. Finally, Yamaguchi (2018) uses using from the Panel Study of Income Dynamics (PSID) and the Dictionary of Occupational Titles (DOT) to study the determinants of the narrowing of the gender wage between 1979 and 1996. He reports a significant increase in women’s cognitive and general skills, which he argues contributed to the reduction in the gender wage gap.

We formalize the mean dynamics hypothesis in our framework as an increase in the mean academic achievement of females relative to males over time which, as we illustrate, can generate the reversal. Under this hypothesis, as under the tail hypothesis, there exists a lower bound of academic achievement $\bar{z}$ such that individuals that attain a given level of schooling $\bar{s}$. Making the normalizing assumption that $\mu_{mt} = 0$ for all $t$, the share of individuals in the cohort that attain a level of schooling of at least $\bar{s}$ at time $t$ is given by:

$$P_{MH}(\mu_{f,t}, \bar{z}_t) \equiv \frac{G_{z_m}(\bar{z}_t) + G_{z_{f,t}}(\mu_{f,t}, \bar{z}_t)}{2}. \quad (6)$$

A change from $E[z_{f,t}] < E[z_m]$ to $E[z_{f,t}] > E[z_m]$ is a necessary and sufficient condition for the gender gap reversal in education to occur, when the distributions of achievement by gender are assumed to have the same variance. Under the mean dynamics hypothesis, the gender ratio among individuals that achieve a minimum level of schooling $\bar{s}$ can be expressed as:

$$R_{MH}^{MH}(\mu_{f,t}, \bar{z}_t) \equiv \frac{G_{z_{f,t}}(\mu_{f,t}, \bar{z}_t)}{G_{z_m}(\bar{z}_t)}, \quad (7)$$

and:

$$\begin{cases} 
R_{MH} < 1 & \text{when } E[z_{f,t}] < E[z_m] \\
R_{MH} = 1 & \text{when } E[z_{f,t}] = E[z_m] \\
R_{MH} > 1 & \text{when } E[z_{f,t}] > E[z_m] 
\end{cases}$$
5 Empirical Estimation

From the Barro-Lee database, we observe the share of individuals that enrolled in or completed a
given level of schooling $s$, in country $i$ and time $t$. We denote it $x_{it}$, which is the empirical measure
of $P(.)$ from our theoretical framework. From the enrollment or completion rates by gender also
available in the data, we can also calculate the female-to-male ratio among those that enrolled
or completed a given level of schooling in country $i$ and time $t$, denoted $y_{it}$. It is the empirical
measure of $R(.)$ in our model. While the original Barro-Lee database consists of 146 countries, we
exclude eight countries that have highly inconsistent or missing time-series data. We also exclude
countries that have a female-to-male ratio in tertiary enrollment or secondary completion above 4,
and countries with age group populations under 10,000 for males or females for any of the years.
These sample restrictions result in a balanced panel of 115 countries which we use as our baseline
estimation sample.

We estimate our model parameters by pooling observations of $x$ and $y$ available every five years
from 1950 to 2010 for each country in our balanced panel. Our baseline pooled specification
therefore exploits variation in educational attainment and gender ratios over time and between
countries. Table III reports the countries with the highest and lowest female-to-male ratio for the
three levels of educational attainment. It shows wide cross-country variation in the gender ratio for
tertiary enrollment and secondary school completion. As of 2010, the 10th-90th percentile range
of the female-to-male ratio in tertiary enrollment spans from 0.59 to 1.73, with a standard deviation
of 0.45. For secondary completion, the 10-90th percentile range of the gender ratio goes from 0.80
to 1.24, with a standard deviation of 0.23. We observe less cross-country variation for the gender
ratio in primary completion, with a 10th-90th percentile range spanning from 0.93 to 1.13, and a
standard deviation of 0.11.

Figure I and Table IV show that relative to cross-country variation, time variation in gender ratios
for tertiary enrollment and secondary completion is also considerable. The average female-to-male
ratio for countries in our sample has increased from about 0.40 in 1950 to about 1.22 in 2010 for
tertiary education, and from 0.55 in 1950 to 1.08 in 2010 for secondary education. The country
with the largest increase in the share of females in tertiary education has been Saudi Arabia, from
0.16 to 2.42 over the period 1950-2010. In contrast with tertiary and secondary education, we observe significantly less variation over time for primary completion, with an increase of the mean female-to-male ratio from 0.82 to 1.03 on average in our baseline estimation sample.

We estimate our model parameters for two alternative levels of educational attainment \( \bar{s} \): enrollment in tertiary education and secondary school completion. In our model, \( P(\cdot) \) and \( R(\cdot) \), the theoretical equivalents of \( x_{it} \) and \( y_{it} \), are jointly determined by \( G_{z_f} \) and \( G_{z_m} \), and by the academic achievement threshold \( \bar{z} \). Unlike \( \bar{z}_t \), \( x_t \) is observable, and the two distributions are assumed to be fully characterized by the four-parameter vector \( \{ \mu_m, \sigma_m^2, \mu_f, \sigma_f^2 \} \). For tertiary education, estimates are obtained by fitting the total enrollment rate in tertiary education and the female-to-male ratio among individuals that enrolled in tertiary education. For secondary school completion, estimates are derived by fitting the secondary school completion rate and the female-to-male ratio among individuals that completed secondary education. The baseline estimation is conducted over the entire period of 1950-2010, with observations available every five years from the Barro-Lee database.

We do not estimate our model for primary school completion for both conceptual and methodological reasons. As highlighted in Section 3, primary school completion has been universal for many countries over the past decades, where primary schooling has been mandatory for a long time. As we illustrated, these features generate limited variation in completion rates and gender ratios over time and across countries to estimate our model parameters. This also implies that primary schooling completion in many countries is the result of institutional features instead of optimal decisions of investment in schooling. This is incompatible with our conceptual framework which assumes that educational attainment is determined by the optimization of individual intertemporal utility functions.

5.1 Tail Hypothesis Parameter Estimation

In this first specification, only the tail hypothesis is assumed to predict the dynamics of the gender gap in educational attainment, while the mean hypothesis plays no role. From our conceptual
framework, the tail hypothesis consists of a total of four parameters \(\{\mu_m, \mu_f, \sigma^2_m, \sigma^2_f\}\). Without any loss of generality, the parameters of the model under the tail hypothesis can be reduced to two, by normalizing one of the two probability density functions. We standardize the female probability density function such that \(f_{z_f} (\bar{z}_t) \sim N(0, 1)\), and therefore estimate the two moments of the male academic achievement distribution \(\{\mu_m, \sigma^2_m\}\).

The model under the tail hypothesis predicts a unique value \(\hat{y}_{it}\), conditional on the triplet \(\{x_{it}, \mu_m, \sigma^2_m\}\). To fit the model under the tail hypothesis, we estimate the two parameters \(\{\mu_m; \sigma^2_m\}\) from our sample of over 100 countries observed in multiple years. From our model, the tail hypothesis after normalization of the academic achievement distribution for females can be expressed empirically as:

\[
y_{it}^{TH} = \frac{G_{z_f}(\bar{z}_{it})}{G_{z_m}(\bar{z}_{it}, \mu_m, \sigma^2_m)} \cdot \exp(\varepsilon_{it}),
\]

where \(\varepsilon_{it} \sim N(0, \sigma^2)\). Taking the logarithm of Equation 8 and substituting \(\bar{z}_{it} = P^{-1}(x_{it}, \mu, \sigma^2)\) yields the final econometric reduced form:

\[
\log y_{it}^{TH} = \log G_{z_f}(P^{-1}(x_{it}, \mu_m, \sigma^2_m)) - \log G_{z_m}(P^{-1}(x_{it}, \mu_m, \sigma^2_m), \mu_m, \sigma^2_m) + \varepsilon_{it}.
\]

Equation 9 is a non-linear mapping from \(x\) to \(y\) defined by the parameters of the males’ academic achievement distribution relative to the females’ distribution. As the error term \(\varepsilon_{it}\) is normally distributed, the model under the tail hypothesis can be fitted numerically by finding the parameter values that minimize the sum of squared errors.

### 5.2 Mean Hypothesis Parameter Estimation

In this second specification, the tail hypothesis is assumed to play no role and only the mean hypothesis predicts the gender gap dynamics in educational attainment. To estimate the parameters of the mean performance hypothesis, we assume that \(\sigma^2_m = 1\) and introduce the term \(t\mu_f\) to allow for an increase in female average performance over time relative to males’. This baseline formulation of the mean hypothesis only allows for a linear time progression of the female mean, but it has the
advantage of making the model under the mean hypothesis more comparable to the tail hypothesis. In this context, one can think of $\mu_m$ as the constant in the model and $t\mu_f$ as the time trend. Thus, $\mu_m - \mu_f$ and $\mu_m - 13\mu_f$ correspond to the relative mean of the males’ distribution compared to the females’ distribution in periods 1 (1950) and 13 (2010), respectively. From equation 7, the structural model under the mean dynamics hypothesis can be expressed as:

$$y_{it}^{MH} = \frac{G_{zf}(z_{it}, t\mu_f)}{G_{zm}(z_{it}, \mu_m)} \cdot \exp(\varepsilon_{it}),$$

(10)

where $\varepsilon_{it} \sim N(0, \sigma^2_{\varepsilon})$.

Taking the logarithm and substituting in a similar fashion as for Equation 9 yields the following reduced form:

$$\log y_{it}^{MH} = \log G_{zf}(P^{-1}(x_{it}, \mu_m, t\mu_f), t\mu_f) - \log G_{zm}(P^{-1}(x_{it}, \mu_m, t\mu_f), \mu_m) + \varepsilon_{it}.$$  

(11)

As for the tail hypothesis, we fit the model numerically to estimate the mean hypothesis two-parameter vector $\{\mu_m, \mu_f\}$ so that the sum of squared errors is minimized.

### 5.3 Joint Estimation

This third specification allows for both the tail and mean hypothesis to jointly predict the dynamics of the gender gap in educational attainment. It allows to quantify the respective explanatory power of each of the two hypotheses, once the influence of the other hypothesis is also accounted for. To test for each of the two hypotheses while controlling for the other hypothesis’ parameter, we jointly estimate the three key parameters of the model $\{\mu_m, \mu_f, \sigma^2_m\}$ using the following equation:

$$\log y_{it}^{BH} = \log G_{zf}(P^{-1}(x_{it}, \mu_m, t\mu_f, \sigma^2_m), t\mu_f) -$$

$$\log G_{zm}(P^{-1}(x_{it}, \mu_m, t\mu_f, \sigma^2_m), \mu_m, \sigma^2_m) + \varepsilon_{it}.$$  

(12)
We also estimate a second version of Equation 12, where we also control for GDP per capita. This allows to check whether our main results are not purely driven by the overall economic development in each country and over time. The data on PPP-adjusted per capita income data are sourced from the Penn World Tables. The drawback of this robustness check is that it has to be conducted in a smaller sample than our baseline sample, as GDP per capita is not available for all years and all countries included in our baseline sample, particularly for early years. As per-capita income data is missing for many countries in years prior to the 1960s, we estimate the results in our GDP sample starting from the 1960s rather than the 1950s.

6 Results

6.1 Baseline results

Table VI reports the parameter estimates of the model for tertiary education enrollment (Upper panel), and for secondary school completion (Lower panel). The estimation is conducted in both the baseline sample with all countries that have complete educational attainment time-series, as well as in a subsample of countries for which full GDP per capita time series are also available. The reduced sample in which GDP per capita is also available for the full period allows to check the robustness of our findings when countries’ changes in levels of economic development over time are also accounted for. For each of the two samples, three different specifications of our theoretical model are estimated: (1) tail hypothesis parameter only (2) mean hypothesis parameter only (3) both mean and tail hypotheses parameters.

When entered individually in separate equations, both the female mean shift per period ($\mu_f$) and the male standard deviation ($\sigma_m$) parameters are statistically significantly different from their null hypothesis values of 0 and 1, respectively. These results hold for both secondary school completion and tertiary enrollment dynamics. For tertiary enrollment, the R-squared of the tail hypothesis and mean dynamics hypothesis specifications reported in column 1 and 2 in the upper panel are very similar. This suggests that both hypotheses have roughly the same explanatory power in accounting
for the dynamics of the gender gap in tertiary enrollment.

Column 3 of Table VI reports the estimation results when both hypotheses’ parameters are jointly estimated. Both the tail and mean hypotheses parameter estimates are individually highly significant for tertiary education. Accounting for both the tail and mean hypotheses parameters increases the R-squared to 0.41 compared to the specification with the tail hypothesis parameters only (0.37) or with the mean hypothesis parameters only (0.36). Accounting for both hypotheses therefore strongly increases the predictive power for tertiary enrollment dynamics.

The lower panel of the table reports our results for secondary education. In contrast with tertiary education estimates, the R-squared of the tail hypothesis is significantly higher (0.49) than that of the mean hypothesis (0.43), for secondary education. The tail hypothesis has thus stronger explanatory power compared to the mean hypothesis for secondary education dynamics. In contrast with tertiary enrollment, the R-squared increases only marginally to 0.50 compared to the model under the tail hypothesis only (0.49), reported in column 1. Additionally, the magnitude of the mean hypothesis parameter $\mu_f$ decreases more than threefold from 0.022 to 0.06 when the tail hypothesis parameter is also accounted for, although it remains statistically significant at the 5% level. These results overall indicate that the mean dynamics hypothesis has little additional explanatory power to the tail hypothesis in accounting for the gender gap dynamics in secondary school completion.

Figure VI displays the fit of our model when the tail and mean hypothesis are separately and jointly estimated. This visual evidence brings additional insights on how the two hypotheses accurately predict the gender gap reversal in educational attainment. The tail hypothesis fitted for tertiary education predicts a relatively late reversal of the gender gap with respect to tertiary enrollment while the mean hypothesis predicts an earlier reversal. When both hypothesis are combined, the reversal is predicted to occur for an even lower value of the total enrollment rate (0.20), which is most consistent with what is empirically observed. According to the Barro-Lee data, the average female-to-male ratio in tertiary enrollment increased reached one when the global enrollment rate in tertiary education was between 0.13 and 0.16. For secondary completion, the mean hypothesis predicts a very early reversal of the gender gap, while the tail hypothesis and the joint hypotheses
predict the reversal to take place for a much higher value of the secondary completion rate of about 0.65. This late reversal is closer to what is observed in the Barro-Lee dataset, where the global gender gap reversal in educational occurs when global secondary completion rate is about 0.50.

6.2 Robustness Checks

Columns 4 to 7 of Table VI display the estimation results from the same regressions as for the full sample, but only for countries where GDP per capita data is available for all years covered by the education data. Compared to our baseline sample, advanced economies are over-represented in the constrained sample, as they are more likely to have complete GDP time-series compared to developing economies.

The results displayed in column 4 to 6 are quite consistent with the estimates in column 1 to 3. For tertiary enrollment, both the mean and tail hypothesis parameters are highly statistically significant when estimated in separate regressions (column 4 and 5) as well as jointly (column 6). As in the baseline specification, the R-squared strongly increases when both hypothesis parameters are included instead of the parameters of just one hypothesis. For secondary school completion, as in the baseline estimation, the R-squared remains unchanged once the tail hypothesis parameter is also accounted for. The tail hypothesis parameter is highly statistically significant, while the mean hypothesis parameter is virtually zero.

In Column 7, GDP per capita is added to our baseline specifications as a control variable. The rationale behind this approach is to check whether our main results are not purely driven by changes in economic development over time that could be associated with both enrollment rates in education and gender ratios. The explanatory power of both the tail and mean hypotheses’ parameters remain statistically significant at the one percent level for tertiary education, and the R-squared is only marginally improved by the inclusion of GDP per capita. Similarly, for secondary education, the R-squared is virtually unchanged by the inclusion of GDP per capita, and remains quite simi-

---

20As GDP per capita is not available for all years and all countries in the Barro-Lee data, the sample size is smaller than in our baseline sample.
lar to that of the specification with the tail hypothesis only reported in column 5. The coefficient on GDP per capita is however positive and statistically significant at the one percent level, and the magnitude of both the tail and mean hypotheses is to some extent reduced, suggesting that our baseline estimates were to some extend capturing the confounding influence of overall economic development. Our main results are however pretty robust to controlling for GDP per capita.

6.3 Magnitude of the parameter estimates

In the constrained sample, which is closer in composition to the 2015 PISA sample that consists of high and middle and high-income countries, estimates for the male-to-female variance ratio - the square of the estimate for the tail hypothesis parameter - vary between 1.14 and 1.32 depending on the specification. The magnitude of the estimates is pretty consistent with the female-to-male variance ratio in test scores estimated by Machin and Pekkarinen (2008) using PISA 2003, as well as from our estimates using a larger sample of countries from PISA 2015, reported in section 4. The average male-to-female variance ratio in test scores in PISA 2015 is 1.21 for reading and 1.20 for mathematics. In addition, the tail hypothesis parameter estimates are very similar when estimated from the secondary school completion or tertiary enrollment time-series. This internal and external consistency of the estimates provides further comfort on the predictive power and validity of the tail hypothesis.

6.4 Interpretation and discussion

The significance of the tail hypothesis parameter estimate $\sigma$ is quite robust across estimation samples and levels of educational attainment. In addition, its magnitude is consistent with the estimates of Machin and Pekkarinen (2008) from international student assessments is relatively stable across specifications. The mean hypothesis parameter $\mu_f$ has strong explanatory power in the tertiary enrollment estimation, even when the tail hypothesis parameter is also accounted for. However, its value is not statistically different from zero for secondary-school completion when both hypotheses are fitted jointly. This indicates that the combined forces of the tail and mean hypotheses have
been strongly at play in tertiary enrollment dynamics, but not in secondary completion dynamics where the mean hypothesis appears to play a weaker role. This finding could help explain why the magnitude of the gender gap reversal in education has been stronger for tertiary enrollment than for secondary school completion. As evidenced in Section , the changed the female-to-male ratio in tertiary enrollment has reached higher levels and increased faster than the gender ratio in secondary school completion.

The contrast between the estimated mean hypothesis parameter $\mu_f$ for secondary school completion and tertiary enrollment is at first glance puzzling. One way to reconcile these findings would be to interpret them through the lens of the mean benefit hypothesis, which, as shown in A3.2, has similar properties as the mean performance hypothesis in our setting. If the net benefits to education have increased disproportionally for women relative to men in tertiary education, but not for secondary education, it is intuitive to observe a strong explanatory power of the mean dynamics hypothesis for tertiary education, but not for secondary education. Those results are consistent with prior contributions that emphasize the role played by the disproportional increase in the female college premium over time relative to males such as Goldin et al. (2006) or Katz and Murphy (1992). In line with this argument, contributions such as Olivetti and Petrongolo (2016) and Ngai and Petrongolo (2017) have argued that secular labor demand shifts favored occupations and industries disproportionately employing female college graduates. This could account for the disproportionate increase in female college premium and hence for the relative increase in women’s tertiary enrollment.

In contrast, the literature is mostly silent on whether a disproportionate increase in the returns to education for females occurred for the secondary level. This is partly because most of the literature on the topic focuses on the US or other advanced economies where secondary school completion has been widespread for some time. However, in many developing economies, secondary school completion is still far from being universal, as illustrated in Table II. In that respect, our findings indirectly shed light on this question by showing that the mean dynamics hypothesis does not appear to explain the gender gap dynamics in secondary school completion level. Our results imply that the disproportionate increase in returns to schooling for women relative to men was non-linear and mostly concentrated at the tertiary level, with little relative change in the returns to education
for females at the secondary level. In contrast, the tail hypothesis parameter significantly predicts the gender gap dynamics for both secondary and tertiary education in all specifications.

7 Conclusion

This paper contributed to a better understanding of the driving forces behind the gender gap reversal in education, both theoretically and empirically. We developed a unifying theoretical framework to account for the gender gap reversal in education, and formulated two main hypotheses for the reversal within our framework. We showed that both the mean and tail dynamics can explain the gender gap reversal in education theoretically. Empirically, we found evidence for both hypotheses having predictive power to account for the gender gap reversal reversal. However, the role played by the tail hypothesis parameter in the gender gap dynamics appears more robust empirically.

Our results indicate that the lower variance in test score achievement among females has been a driver of the observed gender gap reversal in education, a finding that has not been uncovered by previous literature. The larger variability of men’s achievement test scores observed empirically remains mostly unexplained, and could be explored by further research. Our findings also imply that the over-representation of boys at the bottom end of school performance may be of concern for policy makers. Delving into the origins of the over-representation of boys at the bottom of the achievement distribution could help addressing boys’ growing educational disadvantage. Addressing boys’ increasing is important as early school dropouts have been linked to poor labor market performance, higher poverty incidence but also higher crime rates, which men have been shown to be specifically more prone to commit.

This paper suggests that when looking at gender differences in observable outcomes, it is important to go beyond the analysis of means by looking at entire distributions. Fundamentally, when analyzing educational outcomes by gender, the researcher is always looking at truncated distributions. In such distributions, the mean is a function of the dispersion of the underlying distribution. Since evidence shows a higher dispersion of test scores among males, this effect needs to be accounted for, whenever educational achievement is discussed by gender. Our findings also suggest that the
larger variance of test scores among males may be relevant to explain gender differences in other areas. Building on this fact could be a direction for future research in labor economics or other fields of economics.
References


## Tables

Table I: Summary statistics

<table>
<thead>
<tr>
<th></th>
<th>OECD</th>
<th>Non-OECD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Males</td>
<td>Females</td>
</tr>
<tr>
<td><strong>Panel A: Primary completion rate</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1950 Mean</td>
<td>0.70</td>
<td>0.68</td>
</tr>
<tr>
<td>S.d.</td>
<td>0.20</td>
<td>0.21</td>
</tr>
<tr>
<td>2010 Mean</td>
<td>0.98</td>
<td>0.98</td>
</tr>
<tr>
<td>S.d.</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td><strong>Panel B: Secondary completion rate</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1950 Mean</td>
<td>0.22</td>
<td>0.17</td>
</tr>
<tr>
<td>S.d.</td>
<td>0.15</td>
<td>0.14</td>
</tr>
<tr>
<td>2010 Mean</td>
<td>0.79</td>
<td>0.83</td>
</tr>
<tr>
<td>S.d.</td>
<td>0.15</td>
<td>0.13</td>
</tr>
<tr>
<td><strong>Panel C: Tertiary enrollment rate</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1950 Mean</td>
<td>0.06</td>
<td>0.02</td>
</tr>
<tr>
<td>S.d.</td>
<td>0.04</td>
<td>0.03</td>
</tr>
<tr>
<td>2010 Mean</td>
<td>0.34</td>
<td>0.45</td>
</tr>
<tr>
<td>S.d.</td>
<td>0.15</td>
<td>0.15</td>
</tr>
</tbody>
</table>

*Notes.* Primary and secondary completion rates are calculated for individuals age 15-19 and 20-24 respectively, in a given year. Tertiary enrollment rate are calculated for individuals age 25-29. Statistics exclude eight countries with highly inconsistent time series data or missing values.

*Source.* Barro-Lee educational attainment database.
Table II: Share of Countries with a Gender Gap Reversal in Educational Attainment

<table>
<thead>
<tr>
<th>Region</th>
<th>Primary Completion</th>
<th>Secondary Completion</th>
<th>Tertiary Enrollment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1950</td>
<td>2010</td>
<td>1950</td>
</tr>
<tr>
<td>Advanced economies</td>
<td>0.38</td>
<td>0.46</td>
<td>0.21</td>
</tr>
<tr>
<td>East Asia and the Pacific</td>
<td>0.05</td>
<td>0.74</td>
<td>0.00</td>
</tr>
<tr>
<td>Europe and Central Asia</td>
<td>0.50</td>
<td>0.75</td>
<td>0.00</td>
</tr>
<tr>
<td>Latin America and the Caribbean</td>
<td>0.54</td>
<td>0.83</td>
<td>0.12</td>
</tr>
<tr>
<td>Middle East and North Africa</td>
<td>0.06</td>
<td>0.50</td>
<td>0.00</td>
</tr>
<tr>
<td>South Asia</td>
<td>0.00</td>
<td>0.57</td>
<td>0.00</td>
</tr>
<tr>
<td>Sub-Saharan Africa</td>
<td>0.24</td>
<td>0.59</td>
<td>0.06</td>
</tr>
<tr>
<td>OECD</td>
<td>0.42</td>
<td>0.53</td>
<td>0.14</td>
</tr>
<tr>
<td>Non-OECD</td>
<td>0.25</td>
<td>0.63</td>
<td>0.05</td>
</tr>
<tr>
<td>All countries</td>
<td>0.29</td>
<td>0.64</td>
<td>0.06</td>
</tr>
</tbody>
</table>

Notes. Panel A. The gender gap reversal in educational attainment is defined to occur when the enrollment/completion rate of females for a given level of education is higher than that of males. Primary and secondary completion rates are calculated for individuals age 15-19 and 20-24 respectively, in a given year. Tertiary enrollment rates are calculated for individuals age 25-29. Statistics exclude eight countries with highly inconsistent time series data or missing values.

Source. Barro-Lee educational attainment database.
Table III: Female-to-male ratio in educational attainment, Top and Bottom 5 Countries, 2010

<table>
<thead>
<tr>
<th>Highest female-to-male ratio</th>
<th>Lowest female-to-male ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tertiary enrollment</strong></td>
<td></td>
</tr>
<tr>
<td>Saudi Arabia</td>
<td>2.42</td>
</tr>
<tr>
<td>Slovenia</td>
<td>2.29</td>
</tr>
<tr>
<td>Dominican Republic</td>
<td>2.17</td>
</tr>
<tr>
<td>Kazakhstan</td>
<td>2.02</td>
</tr>
<tr>
<td>Uruguay</td>
<td>2.01</td>
</tr>
<tr>
<td><strong>Secondary completion</strong></td>
<td></td>
</tr>
<tr>
<td>Nepal</td>
<td>1.90</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>1.79</td>
</tr>
<tr>
<td>Zimbabwe</td>
<td>1.72</td>
</tr>
<tr>
<td>Sudan</td>
<td>1.58</td>
</tr>
<tr>
<td>Tanzania</td>
<td>1.53</td>
</tr>
<tr>
<td><strong>Primary completion</strong></td>
<td></td>
</tr>
<tr>
<td>Malawi</td>
<td>1.94</td>
</tr>
<tr>
<td>Lesotho</td>
<td>1.85</td>
</tr>
<tr>
<td>Namibia</td>
<td>1.70</td>
</tr>
<tr>
<td>Myanmar</td>
<td>1.53</td>
</tr>
<tr>
<td>Mali</td>
<td>1.43</td>
</tr>
</tbody>
</table>

Notes. The female-to-male ratio is calculated as the enrollment/completion rate of females divided by the enrollment/completion rate of males for a given level of education. Primary and secondary completion rates are calculated for individuals age 15-19 and 20-24 respectively, in a given year. Tertiary enrollment rates are calculated for individuals age 25-29. The ranking excludes eight countries with highly inconsistent time series and missing values, countries with a female-to-male ratio above 4, and countries with a male or female population under 10,000 for the corresponding age group.

Source. Barro-Lee educational attainment database.
Table IV: Female-to-male ratio in educational attainment, Top and Bottom Changers between 1950 and 2010

<table>
<thead>
<tr>
<th></th>
<th>Largest increase in ratio</th>
<th>Largest decrease in ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tertiary enrollment</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saudi Arabia</td>
<td>2.26</td>
<td>Central African Republic</td>
</tr>
<tr>
<td>Slovenia</td>
<td>1.94</td>
<td>Morocco</td>
</tr>
<tr>
<td>Kazakhstan</td>
<td>1.63</td>
<td>Uganda</td>
</tr>
<tr>
<td>Venezuela</td>
<td>1.50</td>
<td>Cambodia</td>
</tr>
<tr>
<td>Croatia</td>
<td>1.47</td>
<td>Niger</td>
</tr>
<tr>
<td><strong>Secondary completion</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nepal</td>
<td>1.84</td>
<td>United Kingdom</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>1.71</td>
<td>Congo</td>
</tr>
<tr>
<td>Tanzania</td>
<td>1.31</td>
<td>Gambia</td>
</tr>
<tr>
<td>Sudan</td>
<td>1.30</td>
<td>Canada</td>
</tr>
<tr>
<td>Tunisia</td>
<td>1.29</td>
<td>USA</td>
</tr>
<tr>
<td><strong>Primary completion</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Malawi</td>
<td>1.69</td>
<td>Mauritania</td>
</tr>
<tr>
<td>Uganda</td>
<td>1.06</td>
<td>Dominican Republic</td>
</tr>
<tr>
<td>Tanzania</td>
<td>1.00</td>
<td>Norway</td>
</tr>
<tr>
<td>Cambodia</td>
<td>0.97</td>
<td>Bostwana</td>
</tr>
<tr>
<td>Egypt</td>
<td>0.90</td>
<td>Russia</td>
</tr>
</tbody>
</table>

*Notes.* The female-to-male ratio is calculated as the enrollment/completion rate of females divided by the enrollment/completion rate of males for a given level of education. Primary and secondary completion rates are calculated for individuals age 15-19 and 20-24 respectively, in a given year. Tertiary enrollment rates are calculated for individuals age 25-29. The ranking excludes eight countries with highly inconsistent time series and missing values, countries with a female-to-male ratio above 4, and countries with a male or female population under 10,000 for the corresponding age group.

*Source.* Barro-Lee educational attainment database.
### Table V: Three hypotheses for the gender gap reversal in education

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Mean of z</th>
<th>Variance of z</th>
<th>$\bar{z}$ (or, equivalently, $b$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tail hypothesis</td>
<td>-</td>
<td>Gender-specific $Var[z_m]$ and $Var[z_f]$</td>
<td>-</td>
</tr>
<tr>
<td>MBH</td>
<td>-</td>
<td>-</td>
<td>Gender-specific $\bar{z}_f$ and $\bar{z}_m$ (or, equivalently, gender-specific $b_f$ and $b_m$)</td>
</tr>
<tr>
<td>MPH</td>
<td>Gender-specific $E[z_f]$ and $E[z_m]$</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

**Notes.** The table reports the assumptions made under the three hypotheses presented in the paper regarding the distribution of academic achievement $z$ and $\bar{z}$, the lower bound of academic achievement such that individuals enroll into a given level of schooling $\bar{s}$.
Table VI: Model Fit of the Tail Hypothesis and the Mean Performance Hypothesis Explaining Sex Ratios in Tertiary and Secondary Education with Enrollment Rates

<table>
<thead>
<tr>
<th></th>
<th>Main sample</th>
<th></th>
<th></th>
<th>Constrained sample</th>
<th></th>
<th></th>
<th>GDP &amp; Both hypotheses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tail hypothesis</td>
<td>Mean performance hypothesis</td>
<td>Both hypotheses</td>
<td>Tail hypothesis</td>
<td>Mean performance hypothesis</td>
<td>Both hypotheses</td>
<td>GDP &amp; Both hypotheses</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
<td>(7)</td>
</tr>
<tr>
<td>Tertiary enrollment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male mean ($\mu_m$)</td>
<td>-0.052** (0.025)</td>
<td>0.497*** (0.010)</td>
<td>0.191*** (0.019)</td>
<td>-0.127*** (0.036)</td>
<td>0.560*** (0.013)</td>
<td>0.171*** (0.022)</td>
<td>0.276*** (0.017)</td>
</tr>
<tr>
<td>Female mean shift per period ($\mu_f$)</td>
<td>0.033*** (0.001)</td>
<td>0.022*** (0.002)</td>
<td>0.040*** (0.002)</td>
<td>0.027*** (0.003)</td>
<td>0.021*** (0.004)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male SD ($\sigma_m$)</td>
<td>1.194*** (0.015)</td>
<td>1.131*** (0.010)</td>
<td>1.232*** (0.020)</td>
<td>1.169*** (0.013)</td>
<td>1.137*** (0.010)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\delta(GDP)$</td>
<td>0.021*** (0.005)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.37</td>
<td>0.36</td>
<td>0.41</td>
<td>0.43</td>
<td>0.40</td>
<td>0.48</td>
<td>0.49</td>
</tr>
<tr>
<td>N</td>
<td>1,469</td>
<td>1,469</td>
<td>1,469</td>
<td>825</td>
<td>825</td>
<td>825</td>
<td>825</td>
</tr>
</tbody>
</table>

Secondary completion

<table>
<thead>
<tr>
<th></th>
<th>Tail hypothesis</th>
<th>Mean performance hypothesis</th>
<th>Both hypotheses</th>
<th>Tail hypothesis</th>
<th>Mean performance hypothesis</th>
<th>Both hypotheses</th>
<th>GDP &amp; Both hypotheses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
<td>(7)</td>
</tr>
<tr>
<td>Male mean ($\mu_m$)</td>
<td>0.109*** (0.019)</td>
<td>0.450*** (0.030)</td>
<td>0.153*** (0.023)</td>
<td>0.070*** (0.022)</td>
<td>0.546*** (0.014)</td>
<td>0.228*** (0.026)</td>
<td>0.371*** (0.019)</td>
</tr>
<tr>
<td>Female mean shift per period ($\mu_f$)</td>
<td>0.022*** (0.005)</td>
<td>0.006*** (0.003)</td>
<td>0.032*** (0.002)</td>
<td>0.016*** (0.003)</td>
<td>0.013*** (0.003)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male SD ($\sigma_m$)</td>
<td>1.173*** (0.013)</td>
<td>1.161*** (0.013)</td>
<td>1.210*** (0.017)</td>
<td>1.174*** (0.012)</td>
<td>1.109*** (0.009)</td>
<td>0.018*** (0.002)</td>
<td></td>
</tr>
<tr>
<td>$\delta(GDP)$</td>
<td>0.49</td>
<td>0.43</td>
<td>0.50</td>
<td>0.51</td>
<td>0.43</td>
<td>0.51</td>
<td>0.52</td>
</tr>
<tr>
<td>N</td>
<td>1,469</td>
<td>1,469</td>
<td>1,469</td>
<td>957</td>
<td>957</td>
<td>957</td>
<td>957</td>
</tr>
</tbody>
</table>

Notes. ***: significant at the 1% level, **: significant at the 5% level, *: significant at the 10% level. Standard errors are bootstrapped with 1000 repetitions and reported in parentheses. TH: tail hypothesis, MPH: mean performance hypothesis. The null hypothesis for $\sigma_m$ is $\sigma_m = 1$. The first three columns are estimated with equations (9), (11) and (12), respectively.
Figures

Figure I: The Gender Gap Reversal in Educational Attainment

Notes. The black line represents the unweighted average for all countries in the sample. Averages are unweighted by country size. In each panel, eight countries were dropped for having highly inconsistent time series or missing observations. The total sample size is 138 countries.

Source. Barro-Lee educational attainment database.
Figure II: The ratio of male variance to female variance in test scores in reading (top panel) and mathematics (bottom panel)

Source. Project for International Student Assessment (PISA) 2015
Figure III: Three Hypotheses for the Gender Gap Reversal in Educational Attainment

Panel A: Tail dynamics hypothesis

Panel B: Mean benefits hypothesis (MBH)

Panel C: Mean performance hypothesis (MPH)

Notes. Panel A: $\bar{z}$ is the threshold of academic achievement above which individuals attain a level of schooling of at least $\bar{s}$, such as tertiary education. The grey arrow indicates a decrease in $\bar{z}$ over time. Panel B: The grey arrows indicate a faster decrease of $\bar{z}_f$ relative to $\bar{z}_m$, the female and male-specific achievement thresholds, over time. The distribution of academic achievement by gender are overlapping as they are identical under the mean benefits hypothesis. Panel C: The arrow pointing to the right indicates an increase in mean academic achievement of females, $\mu_f$, relative to males over time as $\bar{z}$ decreases over time.
Figure IV: Distribution Functions of Academic Achievement by Gender under the Tail Hypothesis: Illustration

Notes. **Panel A:** The two curves show the probability density functions (pdf) of academic achievement, $z$, among males (full line) and females (dashed line), with $\sigma^2_m > \sigma^2_f$ under the tail dynamics hypothesis and $z$ being normally distributed. **Panel B:** The two curves show the complementary cumulative distribution function (ccdf), resulting from the integration from $+\infty$ to $z$ of $f_z(z)$ and $f_{zm}(z)$. The grey arrow indicates a decrease in $\bar{z}$. The gender gap reversal in educational attainment occurs when the ratio of the two CCDFs reaches 1.
Figure V: Relationship between the share of individuals that attain a given level of schooling and the female-to-male ratio under the tail Hypothesis: Illustration

Notes. The x-axis reports the share of individuals that attain a level of schooling of at least $\bar{s}$. The y-axis measures the female-to-male ratio among individuals that attain a level of schooling of at least $\bar{s}$. The gender gap reversal occurs when the female-to-male ratio crosses $y=1$.  

\begin{equation}
1 - F_m(z) + F_f(z) \sqrt{\frac{1}{2}}
\end{equation}
Figure VI: Model Fit: Female-to-male Ratio by level of Educational Attainment

Tertiary enrollment

Secondary completion

Notes. The x-axis measures the enrollment rate in tertiary education for each country-year observation in the region. The y-axis measures the female-to-male ratio among individuals enrolled. The sample sizes are shown in Table ??.
Appendix

A1 Empirical Support for the Model’s Assumptions

A1.1 The Positive Association between Academic achievement and Level of Schooling

The positive relationship between academic achievement and higher levels of schooling is a key assumption of the micro model. We test for the validity of this assumption empirically by using achievement test score data of 10th graders at age 15 for a representative sample of the US population, that was then followed over time beyond high school. The data is from the Educational Longitudinal study (ELS) which started in 2002 when individuals where in 10th grade, and allows to link test scores in 10th grade to whether individuals ever attended post-secondary education in the following years. Using this data, we estimate a linear probability model of Equation (1), where the dependent variable is tertiary education attendance, denoted \( H \), which is a function of the latent variable \( s^* \) in the model:

\[
H(s^*) = \begin{cases} 
1 & \text{if } s^* \geq \bar{s} \\
0 & \text{if } s^* < \bar{s}
\end{cases}
\]

We therefore estimate \( b \) in our model from the binary equivalent of Equation (1) for tertiary education, from the regression:

\[
H_j = \beta z_j + \varepsilon_j
\]

Table A1 reports estimated for the empirical relationship between test scores at age 15 and tertiary education attendance, estimated from the US Educational Longitudinal Survey (ELS). It shows that test scores at age 15, a proxy for individual academic achievement \( z \), is a major determinant of enrollment in tertiary education. The linear probability model coefficient associated with achievement test scores, i.e the estimate for \( b \) in our model, is a positive and highly significant predictor of post-secondary education enrollment a few years later.\(^{21}\)

\(^{21}\)Empirically, the propensity to attend tertiary education increases monotonically with test scores rather than shifting upward discontinuously beyond a given threshold. Our framework is a simplified version of educational enrollment decisions. Empirically, agents certainly enroll into education on the basis of not only individual academic achievement \( z \), but also a set of individual circumstances, such as parental income, personal network or taste for schooling that can be correlated with academic achievement.
Table A1: Model Estimates of the Relationship between Academic achievement and Participation to Tertiary Education, using US Data

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Composite test score</td>
<td>0.449***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading test score</td>
<td>0.449***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mathematics test score</td>
<td>0.450***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of observations</td>
<td>12,538</td>
<td>12,538</td>
<td>12,538</td>
</tr>
</tbody>
</table>

Notes. ***: significant at the 1% level, **: significant at the 5% level, *: significant at the 10% level. Standard errors are reported in parentheses. The regressions are linear probability models run by ordinary least squares. Regressions do not include an intercept. Test scores were standardized to have a mean of 0 and a standard deviation of 1.


A1.2 **The increase in** \( b \) **over time**

A increase in the net benefit of education \( b \), leading to an increase in years of schooling over time, is a key building block of our of our model. To test the validity of this assumption empirically, we estimate a binary version of Equation 1, where the dependent variable is tertiary education attendance, denoted \( H \). \( H \) is observable and is a function of the latent variable \( s^* \) in the model:

\[
H(s^*) = \begin{cases} 
1 & \text{if } s^* \geq \bar{s} \\
0 & \text{if } s^* < \bar{s}
\end{cases}
\]

We estimate \( b \), the net benefits to tertiary education in our model, from the linear probability model:

\[
H_j = \beta z_j + \epsilon_j,
\]

where \( \beta \) denotes the estimate of \( b \) in our model. According to the Barro-Lee data, the gender gap reversal in participation to tertiary education occurred in the mid 1980s in the US. The data used for the estimation is from another longitudinal surveys of 10th graders, the High school and
Beyond (HS&B), which started in 1980. As for the ELS 2002, this study collects 10th grade test scores that can be linked to tertiary education attendance a few years later, allowing to estimate changes in the value of \( b \) over the 20-year period in which the reversal occurred.

Table A2 reports our estimates for \( b \) in the US in 1980 and 2002 using Equation (1) from our framework. It shows that the estimated net benefits of tertiary education increased sharply from 1980 to 2002, providing empirical support to the second building block of our model. The null hypothesis of equality between \( b_{1980} \) and \( b_{2002} \) is strongly rejected by statistical tests. This indicates that individuals with the same level of academic achievement are more likely to enroll in higher education in 2002 than they were in 1980, which is equivalent to higher net benefits of education in 2002 compared to 1980 in our framework.

Table A2: Model Estimates of the Relationship between Academic achievement and Participation to Tertiary Education in 1980 and 2002, using US Data

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{\beta} )</td>
<td>0.278***</td>
<td>0.450***</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>Number of Observations</td>
<td>12,493</td>
<td>12,538</td>
</tr>
</tbody>
</table>

Notes. ***: significant at the 1% level, **: significant at the 5% level, *: significant at the 10% level. Standard errors are reported in parentheses. The regressions are linear probability models run by ordinary least squares. Regressions do not include an intercept. Test score are standardized to have a mean of 0 and a standard deviation of 1 in the full sample.


This evidence is consistent with a large body of literature showing that returns to education, in particular returns to tertiary education, have increased over the past decades. Grogger and Eide (1995) Goldin and Katz (2009) or Acemoglu and Autor (1998) among others provide consistent evidence of a sharp increase of the college wage premium in the US since the beginning of the 1970s.\textsuperscript{22} This rise in the college wage premium was also observed in other contexts. Card and

\textsuperscript{22}The college wage premium is defined as the wage of college-educated workers relative to the wage of high-school educated workers.
Lemieux (2000) also report an important increase in the wage premium of university graduates relative to high school graduates in the UK and Canada over the same period.

A1.3 The Negative Relationship between Enrollment Rate and Mean Academic achievement of those who Enroll

Given the positive association between individual achievement $z$ and optimal level of schooling $s^*$ achieved by individuals in our model, a rise in $b$ leads to a simultaneous increase in enrollment, and to a decrease in $\bar{z}$, the achievement threshold for achieving a given level of schooling $\bar{s}$. We do not observe $\bar{z}$ empirically, but one implication of a decrease in $\bar{z}$ with no accompanying increase in the mean of the distribution is a decrease in the mean ability of individuals enrolled, which can be tested. To do so, we estimate the average verbal test score of individuals that have attended tertiary education in the US over the period 1975–2010, using data from the General Social Survey (GSS). From 1974 onwards, the GSS includes a short 10-item multiple choice test assessing vocabulary knowledge of respondents. A measure of educational attainment in years is also reported, allowing to identify individuals that attended post-secondary education. We classified individuals that have completed more than 12 years of schooling as having attended tertiary education.

Figure A1 depicts the evolution of the average verbal score of post-secondary students relative to the entire population, from 1975 to 2010. In 1975, the average cognitive score of students attending post-secondary education was 0.60 standard deviation higher than the average cognitive score of the whole population. This relative difference decreased progressively until 2005 to reach approximately 0.30 standard deviations in 2010. Consistent with the assumptions of the model, this suggests that greater access to post-secondary education, an increase in $P$ in our framework, was accompanied by a decrease in the average test scores of individuals enrolling in tertiary education.
A2 The Mean Dynamics Hypotheses: Supplementary Material

A2.1 The Mean Benefit Hypothesis (MBH)

Goldin et al. (2006) or Chiappori et al. (2009) suggested that the returns from education, including labor market and marriage market returns, may have risen more for women over the past decades, which could have driven the gender gap reversal. Both contributions invoke the progressive removal of female career barriers, as a driver of increased returns. Goldin et al. (2006) associate the removal of barriers to female employment to a progressive change in fertility and marriage
patterns, driven by the access to reliable contraception (Goldin and Katz (2002)). The authors posit that this raised females’ expectations regarding future labor market outcomes, increased labor force participation, and moved female employment out of traditionally female occupations. In a similar spirit, Chiappori et al. (2009) argue that technological progress has freed women from many domestic tasks, which disproportionally increased the labor market and marriage market returns of schooling for females. According to both contributions, the disproportionate increase in the benefits of education for females relative to males generated the reversal.

We formulate the mean benefits hypothesis in our framework by allowing the net benefits of education \( b \) to differ between genders. Let \( b_m \) and \( b_f \) denote the net benefits of education for males and females, where \( b_f \) and \( b_m \) are allowed to have different dynamics over time. \( G_z(.) \) is assumed to be identical for males and females. Under the MBH, the optimal level of schooling chosen by individuals conditional on academic achievement is gender-specific:

\[
s^* = z \cdot b_g,
\]

or, equivalently, \( \bar{s} \), is gender specific:

\[
\bar{z}_g = \frac{\bar{s}}{b_g}, \tag{14}
\]

where \( g = \{m; f\} \), with \( m \) standing for males and \( f \) for females. Equation 14 states that there exists gender-specific achievement thresholds such that males and females enroll in a given level of schooling \( \bar{s} \), denoted \( \bar{z}_m \) and \( \bar{z}_f \) respectively. The enrollment rate in a given level of schooling \( \bar{s} \) reads:

\[
P^{MBH}(\bar{z}_f, \bar{z}_m) \equiv \frac{G_z(\bar{z}_m) + G_z(\bar{z}_f)}{2}, \tag{15}
\]

Under the MBH, a shift from \( b_m > b_f \) to \( b_m < b_f \) over time, or equivalently, from \( \bar{z}_m < \bar{z}_f \) to \( \bar{z}_m > \bar{z}_f \), is a necessary and sufficient condition for the gender gap reversal in education to occur. The gender ratio among individuals enrolled in a given level of education \( \bar{s} \) is given by:

\[
R^{MBH}(\bar{z}_f, \bar{z}_m) \equiv \frac{G_z(\bar{z}_f)}{G_z(\bar{z}_m)}, \tag{16}
\]
and:

\[
\begin{align*}
R^{MBH} &< 1 \quad \text{if } \bar{z}_m < \bar{z}_f \\
R^{MBH} &= 1 \quad \text{if } \bar{z}_m = \bar{z}_f \\
R^{MBH} &> 1 \quad \text{if } \bar{z}_m > \bar{z}_f.
\end{align*}
\]

[Table ?? about here]

Using the Barro-Lee data, we parameter values of the model by solving the system of two equations (15 and 16) and two unknowns \( \{\bar{z}_{ft}, \bar{z}_{mt}\} \) for each cohort \( t \), where \( P^{MBH}(\bar{z}_{ft}, \bar{z}_{mt}) = y_t \) and \( R^{MBH}(\bar{z}_{ft}, \bar{z}_{mt}) = x_t \).

A3 Technical Appendix

A3.1 The Tail Hypothesis: Mathematical Proofs of Propositions 1 to 3, Normal Distributions

Let \( f_{zf}(z) \) and \( f_{zm}(z) \) denote the probability distribution functions of achievement \( z \) for females and males, respectively. We assume for the sake of the argument that:

\[ z_f \sim N(\mu_f, \sigma_f^2) \]

and

\[ z_m \sim N(\mu_m, \sigma_m^2), \]

where \( \sigma_m^2 > \sigma_f^2 \), according to the tail dynamics hypothesis.

Proof of Proposition 1. The female-to-male ratio \( R^{TH}(\bar{z}) \) tends to zero when the total enrollment rate \( P^{TH}(\bar{z}) \) tends to zero.

First, it is straightforward to see that \( \lim_{\bar{z} \to \infty} P^{TH}(\bar{z}) = \frac{G_{zf}(\bar{z}) + G_{zm}(\bar{z})}{2} = \frac{0+0}{2} = 0 \), where \( G_{\bar{z}}(\bar{z}) \) denotes the complementary cumulative distribution function (or tail distribution function) of achieve-

\[ \text{Estimates are available upon request.} \]

52
ment $z$, defined as $\int_z^{+\infty} f(z) \, dz$.

Let us now study $\lim_{z \to \infty} R^{TH}(\bar{z})$. Using the analytical expression of the probability distribution function of the normal distribution, the ratio $R^{TH}(\bar{z})$ can be expressed as:

$$R^{TH}(\bar{z}) = \frac{\int_{\bar{z}}^{\infty} \frac{1}{\sqrt{2\pi} \sigma_f} e^{-\frac{z^2}{2 \sigma_f^2}} \, dz}{\int_{\bar{z}}^{\infty} \frac{1}{\sqrt{2\pi} \sigma_m} e^{-\frac{(z - \mu_m)^2}{2 \sigma_m^2}} \, dz}$$

Taking the integrals, one can express the ratio as:

$$R^{TH}(\bar{z}) = \frac{\frac{1}{2} \left( 1 - \text{erf} \left[ \frac{\bar{z} - \mu_f}{\sqrt{2 \sigma_f^2}} \right] \right)}{\frac{1}{2} \left( 1 - \text{erf} \left[ \frac{\bar{z} - \mu_m}{\sqrt{2 \sigma_m^2}} \right] \right)},$$

where $\text{erf}(\cdot)$ denotes the Gauss error function, defined as $\text{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} \, dt$.

Using the analytical expression of $R^{TH}(\bar{z})$, we get:

$$\lim_{\bar{z} \to \infty} \frac{\frac{1}{2} \left( 1 - \text{erf} \left[ \frac{\bar{z} - \mu_f}{\sqrt{2 \sigma_f^2}} \right] \right)}{\frac{1}{2} \left( 1 - \text{erf} \left[ \frac{\bar{z} - \mu_m}{\sqrt{2 \sigma_m^2}} \right] \right)} = \lim_{\bar{z} \to \infty} \frac{1 - (\text{erf} \left[ \bar{z} \right])}{1 - (\text{erf} \left[ \bar{z} \right])} = \frac{1 - 1}{1 - 1} = 0,$$

where the second to last step follows from the fact that $\lim_{\bar{z} \to \infty} \text{erf}(\bar{z}) = 1$.

Using the l’Hôpital rule, we take the derivative of the denominator and the numerator to get the following expression:

$$\lim_{\bar{z} \to \infty} \frac{\sigma_m^2}{\sigma_f^2} \exp \left\{ \frac{(\bar{z} - \mu_m)^2}{\sigma_m^2} - \frac{(\bar{z} - \mu_f)^2}{\sigma_f^2} \right\} = \lim_{\bar{z} \to \infty} \frac{\sigma_m^2}{\sigma_f^2} \exp \left\{ \frac{(\bar{z} - \mu_m)^2 \sigma_f^2}{\sigma_m^2 \sigma_f^2} - \frac{(\bar{z} - \mu_f)^2 \sigma_m^2}{\sigma_f^2 \sigma_m^2} \right\}$$

$$= \lim_{\bar{z} \to \infty} \frac{\sigma_m^2}{\sigma_f^2} \exp \left\{ \frac{\bar{z}^2 \sigma_f^2 - 2\bar{z} \mu_m \sigma_f^2 + \mu_m^2 \sigma_f^2 - \bar{z}^2 \sigma_m^2 + 2\bar{z} \mu_f \sigma_m^2 - \mu_f^2 \sigma_m^2}{\sigma_m^2 \sigma_f^2} \right\}$$

53
\[
\lim_{z \to \infty} \frac{\sigma_m^2}{\sigma_f^2} \exp \left\{ \frac{z}{\sigma_m^2 \sigma_f^2} \left[ \sigma_f^2 \sigma_m^2 + 2 \mu_f \sigma_f^2 - 2 \mu_m \sigma_m^2 \right] \right\} = 0,
\]

since, by assumption under the tail dynamics hypothesis, \(\sigma_m^2 > \sigma_f^2\), which are both positive by definition.

**Proof of Proposition 2.** The female-to-male ratio \(R^{TH}(\bar{z})\) tends to one when the total enrollment rate \(P^{TH}(\bar{z})\) tends to one.

First, it is straightforward to see that \(\lim_{\bar{z} \to \infty} P^{TH}(\bar{z}) = \frac{G_{sf}(\bar{z}) + G_{sm}(\bar{z})}{2} = \frac{1}{2} = 1\).

Let us now study the behavior of \(R^{TH}(\bar{z})\) when \(\bar{z}\) tends to \(-\infty\):

\[
\lim_{\bar{z} \to -\infty} \frac{1}{2} \left( 1 - \text{erf} \left[ \frac{\bar{z} - \mu_f}{\sqrt{2} \sigma_f} \right] \right) = \lim_{\bar{z} \to -\infty} \frac{1}{1 - \text{erf} \left[ \bar{z} \right]} = \frac{1 + 1}{1 + 1} = 1,
\]

where we use the fact that \(\lim_{\bar{z} \to -\infty} \text{erf} \left[ \bar{z} \right] = -1\).

**Proof of Proposition 3.** There exists a value of \(P^{TH}(\bar{z})\) such that \(R^{TH}(\bar{z}) = 1\). This value is unique and always exists.

Let us now show that given our distributional assumptions, there exists a value of \(z\) denoted \(z^*\), such that the numerator and denominator are of equal value, thus the ratio is one. Again, we invoke the ratio:

\[
R^{TH}(\bar{z}) = \frac{1}{2} \left( 1 - \text{erf} \left[ \frac{\bar{z} - \mu_f}{\sqrt{2} \sigma_f} \right] \right).
\]

Since we know that the error function is monotonously increasing on the whole domain, \(R^{TH}(\bar{z}) = 1\) when:

\[
\frac{\bar{z} - \mu_f}{\sqrt{2} \sigma_f} = 1 \Leftrightarrow \frac{\bar{z} - \mu_f}{\sigma_f^2} = \frac{\bar{z} - \mu_m}{\sigma_m^2} \Rightarrow \bar{z} = \frac{\mu_m \sigma_f^2 - \mu_f \sigma_m^2}{\sigma_f^2 - \sigma_m^2}.
\]

This equation has a unique solution given \(\sigma_m^2 > \sigma_f^2\). Since the support of \(E(\bar{z})\) is the whole real
line, there always exists a value of $\bar{z}$ denoted $\bar{z}^*$ such that:

$$\bar{z}^* = \frac{\mu_m \sigma_f^2 - \mu_f \sigma_m^2}{\sigma_f^2 - \sigma_m^2}.$$ 

In addition, $\bar{z}^*$ is unique given the vector of exogenous parameters $\{\mu_f, \mu_m, \sigma_f^2, \sigma_m^2]\}.

### A3.2 Equivalence of the MPH and MBH hypotheses for uniform distribution within the intersection of their domains

The MPH in our framework models the gender gap reversal such that females initially have lower participation rates in education, due to a lower $E[z]$. Conditional on $z$, the net benefits of schooling and the optional level of schooling chosen are identical for both genders: $E[s^*_f|z] = E[s^*_m|z]$. The MPH and MBH thus differ in the interpretation of the source of the dynamics in the gender mean differences. They are, in the way they are modeled in our framework, algebraically very similar. It can be shown that if academic achievement is normally distributed, they are actually identical within the intersection of the domains of the two distributions. As a result, the closer the CCDFs of $z$ are to linear, the more similar the two hypotheses will be in our framework. Here follows the proof to the claim that the MPH is equivalent to the MBH for uniform distribution within the intersection of their domains.

**MBH:** For each time period $t$, let $b \equiv a + c$ and $z \sim unif(a, b) = unif(a, a + c)$. Then, $G_z(\bar{z}) = 1 - \frac{\bar{z} - a}{b-a} = 1 - \frac{\bar{z} - a}{a+c - a} = 1 - \frac{\bar{z} - a}{c}$. Also, let (*) $\bar{z}_m \equiv \bar{z}$ and (**): $\bar{z}_f \equiv \bar{z} + a - a_f$. Now, for $\bar{z}_f, \bar{z}_m \in \{a, a + c\}$ $\iff \bar{z} + a - a_f, \bar{z} \in \{a, a + c\}$,

$$E^{MBH}(\bar{z}_f, \bar{z}_m) = \frac{G_z(\bar{z}_m) + G_z(\bar{z}_f)}{2} = \frac{1}{2} \left( 1 - \frac{\bar{z}_m - a}{(a + c) - a} + 1 - \frac{\bar{z}_f - a}{a + c - a} \right) = \frac{1}{2} \left( \frac{\bar{z}_m + \bar{z}_f - 2a}{2c} \right)$$

$$= 1 - \frac{\bar{z}_m + \bar{z}_f - 2a}{4c} \equiv 1 - \frac{2\bar{z} - a - a_f}{4c}$$ and
\[ R^{MBH}(\bar{z}_f, \bar{z}_m) \equiv \frac{G_z(\bar{z}_f)}{G_z(\bar{z}_m)} = 1 - \frac{\bar{z}_f - a}{(a + c) - a} / 1 - \frac{\bar{z}_m - a}{(a + c) - a} = \frac{c - \bar{z}_f - a}{c} / c - \frac{\bar{z}_m - a}{c} \]

\[ = \frac{c - \bar{z}_f + a}{c - \bar{z}_m + a} = \frac{c - \bar{z} + a_f}{c - \bar{z} + a} \]

**MPH:** For \( \bar{z} + a - a_f, \bar{z} \in \{a, a + c\} \),

\[ E^{MPH}(\mu_f, \bar{z}) = \frac{G_{zm}(\bar{z}) + G_{zf,t}(\mu_f, \bar{z})}{2} = \frac{1}{2} \left( 1 - \frac{\bar{z} - a}{(a + c) - a} + 1 - \frac{\bar{z} - a_f}{(a_f + c) - a_f} \right) \]

\[ = \frac{1}{2} \left( 2 - \frac{2\bar{z} - a - a_f}{2c} \right) = 1 - \frac{2\bar{z} - a - a_f}{4c} = E^{MBH}(\bar{z}_f, \bar{z}_m) \text{ and} \]

\[ R^{MPH}(\mu_f, \bar{z}) \equiv \frac{G_{zf,t}(\mu_f, \bar{z})}{G_{zm}(\bar{z})} = 1 - \frac{\bar{z} - a_f}{(a_f + c) - a_f} / 1 - \frac{\bar{z} - a}{(a + c) - a} = \frac{c - \bar{z} - a_f}{c} / c - \frac{\bar{z} - a}{c} \]

\[ = \frac{c - \bar{z} + a_f}{c - \bar{z} + a} = R^{MBH}(\bar{z}_f, \bar{z}_m) \]

Thus, the MBH and MPH are equivalent when achievement is uniformly distributed and when \( \bar{z} + a - a_f, \bar{z} \in \{a, a + c\} \). The difference between the two hypotheses originates from the non-linearities in \( G_z(\cdot) \).