Distributional Growth Accounting: Education and the Reduction of Global Poverty, 1980-2022

Amory Gethin*

Job Market Paper

October 13, 2023

Click here for the most recent version

Abstract

This article studies the role played by education in the decline of global poverty. Drawing on a simple model of education and the wage structure, I propose tools for "distributional growth accounting," extending growth decomposition analysis to allow gains from schooling to vary by income group. I bring this framework to the data by combining a new micro-database representative of nearly all of the world's population, new estimates of the private returns to schooling, and historical education and income distribution statistics. Under conservative assumptions, education accounts for 50% of global economic growth, 70% of income gains among the world's poorest 20% individuals, and 40% of extreme poverty reduction since 1980. It also explains over 50% of improvements in the share of labor income accruing to women. Combining these findings with measures of direct government redistribution from a companion paper brings the contribution of public policies to global poverty reduction to at least 50%.

*Paris School of Economics (Grant ANR-17-EURE-001). I am particularly grateful to Thomas Piketty, Vincent Pons, Emmanuel Saez, and Gabriel Zucman for their continuous guidance. I thank Pierre Bachas, Denis Cogneau, Eva Davoine, Jérémy Do Nascimento Miguel, Sebastian Hager, Eric Hanushek, Hilary Hoynes, Anders Jensen, Pete Klenow, Giulia La Mattina, Ségal Le Guern Herry, Wouter Leenders, David Levine, Antoine Levy, Hongbin Li, Rosanne Logeart, Nicolas Longuet Marx, Edward Miguel, Mathilde Muñoz, Emy Nakamura, Jon Piqueras, Tom Raster, Anne-Sophie Robilliard, Benjamin Schoefer, Andrei Shleifer, Jón Steinsson, Christopher Walters, Danny Yagan, and Céline Zipfel for useful comments. This paper benefited from fruitful discussions with seminar participants at UC Berkeley and the Paris School of Economics. I also thank the International Labor Organization for giving me access to their collection of harmonized household surveys. Juan Ferrer provided excellent research assistance.

1. Introduction

The past decades witnessed dramatic improvements in access to basic public services among the global poor. These improvements were reflected in major progress on indicators as diverse as school enrollment, literacy, vaccination rates, and access to drinkable water (United Nations, 2023). Valued as a transfer received by households, the direct consumption of public goods can account for about 20% of global poverty reduction since 1980. Total government redistribution, including cash and in-kind transfers, can account for 30% (Gethin, 2023).

How useful these policies have been at generating pretax income growth remains, however, an open question. Education, in particular, is often thought to be a productive investment, yet its contribution to global poverty reduction is uncertain. Expanding education affects both economic growth and its distribution within each country. These effects will depend on who benefits from educational expansion, the associated returns to schooling, and general equilibrium effects. Because of difficulties at quantifying these different channels, how large benefits from schooling have been for the global poor remains unsettled. Answering this question holds critical importance for policy, in a world where the vast majority of children from low-income households are enrolled in public schools.

This article makes a first attempt at estimating the aggregate and distributional effects of worldwide educational expansion since 1980. The starting point is a new micro-database representative of 95% of the world's population, which I assemble from various repositories of household surveys and country-specific sources. The data cover individual labor income, education, age, gender, and other socioeconomic variables for 150 countries in 2019, providing a unique snapshot on the contemporary structure of global poverty and inequality.

Starting from this new dataset, I estimate what the world distribution of income would look like if there had been no improvement in schooling since 1980. To construct this counterfactual, I combine standard growth accounting tools with a simple model of education and the wage structure *à la* Goldin and Katz (2007). In this "distributional growth accounting" framework, expanding education increases aggregate labor income by the private return to schooling, which is endogenous to the relative supply of skilled workers in the economy. Educational expansion also has distributional effects by pushing down the relative wage of skilled workers as their relative supply increases. The magnitude of these effects is governed by a long-run elasticity of substitution between skill groups, which I calibrate from the recent macroeconomics literature.

I bring this framework to the data by constructing a counterfactual world distribution of income in four steps. First, I downgrade education levels in each survey until matching the distribution of educational attainment that prevailed in 1980. Second, I reduce individual incomes accordingly, using new measures of the Mincerian returns to primary, secondary, and tertiary education directly estimated from the microdata in each country. Third, I adjust relative wages: the supply of skilled workers would be lower, and hence their relative wage higher, if education had not improved. Finally, comparing the resulting counterfactual to the actual evolution of pretax incomes yields an estimate of the contribution of education to economic growth for different groups within each country. This approach is analogous to the canonical growth accounting exercise, which typically combines cross-country data on workers' average years of schooling with a uniform 10% Mincerian return to derive the same counterfactual (e.g., Barro and Lee, 2015). The main contribution of this paper lies in the use of rich survey microdata representative of 95% of the world's population with a model embedding imperfect substitution between skilled and unskilled workers. Together, these two ingredients allow for a much more granular estimation of how the economic benefits of education vary within and across countries.

I validate this approach with new quasi-experimental evidence from three large-scale schooling initiatives in India, Indonesia, and the United States. Combining data on the distribution of income with differential exposure to each program across subnational regions, I estimate large causal effects of education on regional incomes, comparable to individual returns found in the same contexts. The three policies also disproportionately benefited low-income earners, generating large reductions in inequality. The distributional growth accounting framework reproduces these two findings with a remarkable degree of accuracy, suggesting that it provides a good methodological foundation to study the role of education in global poverty reduction.

In my benchmark specification, I find that private returns to schooling can account for about 50% of global economic growth and 70% of income gains for the world's poorest 20% individuals since 1980 (see figure 1). They also explain about 40% of the reduction in the share of the world's population living in extreme poverty. Given the predominant role that governments have had at providing education and other basic services to low-income households, this puts public policies at the center of the historical fall of global poverty. Combining measures of direct government redistribution from a companion paper (Gethin, 2023) with indirect investment benefits from education estimated in this paper brings the total contribution of public policies to global poverty reduction to at least 50%.

These estimates should be considered as conservative. The estimation relies on standard Mincerian returns to schooling, which are typically lower than causal estimates derived from natural experiments. It is based on a relatively high elasticity of substitution between skill groups, limiting redistributive effects generated by the growing relative supply of skilled workers. It assumes that education only affects labor income, ignoring potential effects on capital income and productivity through savings, innovation, or other channels (Gennaioli et al., 2013; Queiró, 2022). It also ignores human capital externalities, on which there is now significant empirical evidence.[1] All in all, my findings are governed by two main sets of parameters: the private returns to schooling and the degree of imperfect substitutability between skill groups. With plausible values for these parameters, I bound the contribution of education to the world's poorest 20% growth between 60% and 90%. Moving below 60% would require assuming either that workers are perfect substitutes, or that returns to schooling are significantly below those found in the data, in contradiction with much of the labor economics literature and my own analysis of the three natural experiments mentioned above.

Methodologically, accounting for distributional effects within countries appears to be crucial to adequately measure the role of schooling in global poverty reduction. A standard growth decomposition relying on cross-country data, as in Barro and Lee (2015), would underestimate the contribution of education by a factor of three. One reason is simply that cross-country data cannot accurately measure poverty: the poorest individuals in the world are not the poorest countries. More importantly, the classic approach fails to account for important dimensions of educational expansion, such as the greater concentration of labor incomes at the bottom of the distribution and general equilibrium effects redistributing schooling gains from high-skilled to low-skilled workers. That being said, relying on microdata instead of aggregate data also affects some findings in the opposite direction. Moving from a constant Mincerian return of 10%, as often assumed in the literature, to heterogeneous returns by level reduces the contribution of education to global poverty reduction by about 20%. The main reason is that the return to basic education is particularly low in developing countries, ranging from just 3% per year in India to 6% in Sub-Saharan Africa.

One should also stress that the large contribution of schooling to global poverty reduction does not preclude that other factors, such as physical capital or technology, may have played a significant role too. In the model, for instance, the return to schooling increases with the skill bias of technology. Had technology not improved, the return to schooling would likely be substantially lower than the one observed today. Skill-biased technical change has thus potentially played a key role in *amplifying* the economic benefits of schooling. Quantifying this exact contribution would necessitate survey data going back to the 1980s for all countries in my sample, which are unfortunately not available. For a subset of countries with historical survey data, I provide suggestive evidence that skill-biased technical change has had such positive

[1] Existing studies have generally found strong indications of externalities from higher education. Evidence on other levels of schooling is more debated. See in particular Acemoglu and Angrist (2000), Chauvin et al. (2017), Ciccone and Peri (2006), Glaeser and Lu (2018), Guo, Roys, and Seshadri (2018), Iranzo and Peri (2009), Moretti (2004), and Wantchekon, Klašnja, and Novta (2015).

effects, typically enhancing the benefits of schooling by 20-30%.

Finally, I extend distributional growth accounting to the study of another major historical transformation: the decline of global gender inequality. To do so, I quantify how large gender labor income gaps would be today if education had not improved since 1980. The estimation accounts for differential educational expansion by gender, but also for heterogeneous effects of schooling on earnings and labor force participation. This counterfactual is then compared to the actual evolution of female labor income shares, on which data is available since 1991. The main conclusion is that education can explain a large share of reductions in gender inequality observed in the past decades, typically 50% to 80% depending on the specification and world region considered. Education has thus played a key role in the historical empowerment of women observed in most parts of the world.

A large literature in labor economics uses the canonical labor supply-and-demand framework to relate changes in the wage distribution to educational expansion.[2] Concurrently, a considerable literature in macroeconomics investigates the contribution of human capital to development and economic growth.[3] These two methodological perspectives, one focused on within-country inequality and the other on cross-country dynamics, have remained relatively independent from one another. The main contribution of this article is to bring them together into a unified "distributional growth accounting" framework, which I use to quantify the role of education in the reduction of global poverty and gender inequality since 1980.

This article also contributes to our understanding of the forces shaping the long-run evolution of the world distribution of income. Global inequalities have undergone profound transformations in recent decades, including rapidly declining poverty (Chen and Ravallion, 2010; Pinkovskiy and Sala-i-Martin, 2016; Sala-i-Martin, 2006), the emergence of a new "global median class" (Lakner and Milanovic, 2016), skyrocketing top income inequality (Chancel and Piketty, 2021), and moderately decreasing gender inequality (Neef and Robilliard, 2021). Amongst the numerous factors shaping these dynamics, I isolate the contribution of one of them: education. I

[2] This framework has been used extensively to account for trends in wage inequality in the United States (Acemoglu and Autor, 2011; Autor, Goldin, and Katz, 2020; Deming, 2023; Goldin and Katz, 2007; 2008; Hershbein, Kearney, and Pardue, 2020; Katz and Murphy, 1992). A growing literature successfully extends this analysis to low- and middle-income countries: see for instance Fernández and Messina (2018), Khanna (2023), and Vu and Vu-Thanh (2022). A few studies also investigate the role of education in explaining cross-country differences in gender inequality (e.g., Kleven and Landais, 2017).

[3] The recent literature has focused more heavily on explaining cross-country differences in development: see for instance Caselli and Coleman (2006), Gennaioli et al. (2013), Hall and Jones (1999), Hanushek and Kimko (2000), Hendricks and Schoellman (2018), Hsieh and Klenow (2010), Jones (2014), and Rossi (2020). Growth accounting decompositions have also been used extensively since Solow (1957), although worldwide perspectives are more recent (e.g., Barro and Lee, 2015; Mankiw, Romer, and Weil, 1992). Most closely related to this paper is recent work by Collin and Weil (2020), who estimate that accelerating human capital accumulation could have significant effects on global poverty reduction in coming decades.

find that it has been a powerful source of convergence and can account for a large share of real income gains for both women and the world's poorest individuals.

Finally, this paper relates to the large empirical evidence on the economic effects of education. A vast literature documents positive impacts of education on individual earnings (Card, 2001; Deming, 2022). A more limited number of studies have gone beyond individual outcomes to study the general equilibrium effects of education policies (e.g., Che and Zhang, 2018; Duflo, 2004; Khanna, 2023). I contribute to this body of work by revisiting previously studied natural experiments in India (Khanna, 2023), Indonesia (Duflo, 2001), and the United States (Acemoglu and Angrist, 2000) to shed light on the aggregate and distributional effects of educational expansion. I also draw heavily on existing studies to calibrate the parameters guiding my results, such as elasticities of substitution between skill groups and differential economic effects of education by gender. In doing so, this article is an attempt at taking the best of microeconomic evidence to draw conclusions on the aggregate effects of education on global poverty and gender inequality.

The rest of the paper is organized as follows. Section 2 outlines the conceptual framework used for distributional growth accounting. Section 3 presents the data and methodology. Section 4 describes the main results on the role of education in shaping the distribution of global economic growth. Section 5 turns to the study of global gender inequality. Section 6 provides a general discussion and additional results. Section 7 concludes.

2. Distributional Growth Accounting

This section presents the framework used to estimate the aggregate and distributional effects of human capital accumulation. Section 2.1 formulates the problem of interest. Sections 2.2 and 2.3 expose simple formulas relating the distribution of educational attainment to aggregate earnings and inequality. Section 2.4 outlines the methodology used to estimate the contribution of educational expansion to the distribution of economic growth.

2.1. Setup

2.1.1. Research Question

Output is produced by combining physical capital K and workers with different levels of educational attainment:

$$Y = F(K, L) = F(K, L_0, L_1, ..., L_m) \tag{1}$$

Throughout the paper, I define workers as including both wage earners and the self-employed. Labor income includes both compensation of employees and mixed income, referred to as "wages" for simplicity. Capital income includes all remaining national income components.

Workers of skill s are paid their marginal product w_s. Skill groups may be imperfectly substitutable in production, implying that wages depend on relative supplies:

$$w_s = w_s(L) \tag{2}$$

Consider a group p, receiving income from both labor and capital, and composed of individuals with different levels of skills. This can correspond to a group of the income distribution, such as the poorest 20%, or to a social group such as women. The average income of group p is the sum of their capital income and labor income:

$$y^p = y_K^p + \sum_s w_s(L)\, L_s^p \tag{3}$$

With L_s^p the share of workers of type s in group p. I make the conservative assumption that capital income is not affected by schooling, as in standard growth accounting (Barro and Lee, 2015). We are interested in estimating a counterfactual income $\tilde{y}^p$, given a counterfactual distribution of educational attainment $\tilde{L} = (\tilde{L}_1, ..., \tilde{L}_m)$:

$$\tilde{y}^p = y_K^p + \sum_s w_s(\tilde{L})\, \tilde{L}_s^p \tag{4}$$

To estimate $\tilde{y}^p$, we therefore need to characterize four sets of parameters: the initial joint distribution of labor and capital incomes; the initial joint distribution of wages $w_s(L)$ and schooling L; counterfactual education levels $\tilde{L}_s^p$; and counterfactual wages $w_s(\tilde{L})$.

2.1.2. Model Specification

Throughout the paper, I consider variants of the CES production function with two skill groups:

$$Y = \left(A_H L_H^{\frac{\sigma-1}{\sigma}} + A_L L_L^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \tag{5}$$

With A_H and A_L labor-augmenting technology terms, L_H and L_L the supplies of high-skill and low-skill labor, and σ the elasticity of substitution between H and L.

Assuming that workers are paid their marginal product, the return to schooling is:

$$r(L) = \log\left(\frac{w_H}{w_L}\right) = \frac{\sigma-1}{\sigma}\log\left(\frac{A_H}{A_L}\right) - \frac{1}{\sigma}\log\left(\frac{L_H}{L_L}\right) \tag{6}$$

The relative wage of skilled workers depends on their relative supply, as well as on the skill bias of technology. In particular, a 1% increase in the relative share of skilled workers is associated with a $\frac{1}{\sigma}$% decline in the skill premium. A higher elasticity of substitution implies greater substitutability between skill groups and a lower sensitivity of wages to relative supplies.

2.2. Aggregate Returns to Schooling

2.2.1. Individual Returns to Schooling and Supply Effects

I now characterize the aggregate effects of educational expansion in this model. We are interested in the effect of increasing L_H from an initial level L_H to a new level $\tilde{L}_H$. Denote the resulting change in the supply of skilled workers as $\Delta L_H = \tilde{L}_H - L_H$ and corresponding changes in wages as Δw_L and Δw_H. The effect of skill upgrading on output is:

$$\Delta Y = \tilde{Y} - Y = \Delta L_H \Bigg(\underbrace{r(L)}_{\substack{\text{Initial Returns}\\ \text{to Schooling}\\ >0}} + \underbrace{\Delta r(L)}_{\substack{\text{Supply Effects}\\ \text{on Newly Skilled}\\ <0}} \Bigg) + \underbrace{L_H \Delta w_H}_{\substack{\text{Supply Effects}\\ \text{on Always Skilled}\\ <0}} + \underbrace{L_L \Delta w_L}_{\substack{\text{Supply Effects}\\ \text{on Unskilled}\\ >0}} \tag{7}$$

With $r(L) = w_H - w_L$ the return to schooling at baseline and $\Delta r(L) = r(\tilde{L}) - r(L)$ the change in return to schooling induced by skill upgrading. The effect of human capital accumulation on earnings can be separated into four parts. First, educational upgrading increases the earnings of newly skilled workers through returns to schooling observed at baseline (initial returns to schooling). However, the resulting increase in the supply of skilled workers exerts downward pressure on these returns, mitigating the final benefits for this group (supply effects on newly skilled). For the same reasons, it also reduces the earnings of workers who were already skilled (supply effects on always skilled). Finally, low-skilled workers see their wage increase: the decline in their relative supply drives up their marginal productivity (supply effects on unskilled).

Supply effects thus mitigate the returns to schooling for those benefiting from educational expansion, at the same time as they redistribute income from high-skilled to low-skilled workers. In the case in which low-skilled and high-skilled workers are perfect substitutes, supply effects boil down to zero, and the change in output becomes $\Delta Y = \Delta L_H \, r(L) = \Delta L_H \, r$.

2.2.2. Initial, Final, and True Returns to Schooling

How do these four effects play out in practice, and where does the true aggregate effect of schooling lie? There are two natural options.

Initial Return One option would be to use the initial individual return to schooling observed *before* educational expansion, $r(L)$. This amounts to assuming that supply effects redistribute income between skill groups, but do not affect aggregate output. For instance, if the individual return declined from 10% to 8%, then 10% is the return that should be used. The change in output is then:

$$\Delta \bar{Y} = \Delta L_H r(L)$$

Final Return An alternative option would be to use the final return to schooling observed *after* educational expansion, $r(\tilde{L}) = r(L) - \frac{1}{\sigma}\Delta L_H$. This amounts to assuming that all supply effects are a net loss for the economy. If the individual return declined from 10% to 8%, then 8% is the return that should be used:

$$\Delta \underline{Y} = \Delta L_H r(\tilde{L})$$

True Return The true effect turns out to lie in-between. The aggregate effect of education on output is lower than the initial individual return to schooling. Indeed, with imperfect substitution, there are decreasing returns to human capital accumulation: positive supply effects on unskilled workers are more than offset by negative supply effects on skilled workers (Caselli and Ciccone, 2013). Yet, because of decreasing returns, the true effect is also higher than the final return. Newly skilled workers end up with lower benefits than they might have hoped, but part of this loss benefits, on net, the unskilled:

$$\underbrace{r(\tilde{L})}_{\text{Final Returns}} \leq \underbrace{\frac{\Delta Y}{\Delta L_H}}_{\substack{\text{Aggregate Gains} \\ \text{Per Newly Skilled}}} \leq \underbrace{r(L)}_{\text{Initial Returns}} \tag{8}$$

The interested reader will find a graphical illustration providing the main intuition in appendix figure B1. In the context of this paper, one is interested in estimating the effect of reducing education back to its 1980 level. One could use the 2019 (final) return to schooling, corresponding to the derivative of the production function before reducing education. Alternatively, one could use marginal gains from schooling observed after reducing education, corresponding to counterfactual (initial) returns. With imperfect substitution, the log of output is concave in schooling, implying that the true effect of education lies in-between these two estimates. As the elasticity of substitution increases, the production function becomes less concave: the returns to schooling observed in 2019 become a better approximation of the output loss that would result from reducing education.

Assuming that the parameters of the production function are known, it is possible to re-express the true change in output as a function of a "true individual return" r^* that should be used. This return satisfies:

$$\tilde{Y} = w_H L_H + w_L \tilde{L}_L + \exp\left(\log(w_L) + r^*\right)\Delta L_H \tag{9}$$

Put simply, new output is the sum of wages received by the always skilled L_H at baseline (first term), wages received by the always unskilled $\tilde{L}_L$ at baseline (second term), and wages received by the newly skilled ΔL_H (third term), whose educational upgrading increases output by r^* log points. Rearranging yields a closed-form solution for the true aggregate return to schooling:

$$r^* = \log\left(\frac{\tilde{Y} - w_H L_H - w_L L_L}{\Delta L_H}\right) - \log(w_L) \tag{10}$$

Appendix B.2 provides a theoretical discussion, as well as results from a simple simulation using a CES production function, illustrating how the optimal return to schooling differs from initial and final returns depending on the elasticity of substitution and the skill bias of technology. For parametrizations similar to those found in the data, the optimal return to schooling is a weighted average of initial and final returns, with a typical weight on initial returns of 50-70%.

2.3. Distributional Effects of Schooling

I now turn to the effect of educational expansion on the income distribution. Consider two groups, rich R and poor P, who differ in their relative proportions of skilled and unskilled workers: $L_H^R > L_H^P$. Using equation 7, the effect of educational expansion on the rich-poor income gap can be expressed as:

$$\Delta Y^R - \Delta Y^P = \underbrace{\left(L_H^R - L_H^P\right)\Delta w_H}_{\substack{\text{Differential Supply} \\ \text{Effects on Always Skilled} \\ \leq 0}} + \underbrace{\left(L_L^R - L_L^P\right)\Delta w_L}_{\substack{\text{Differential Supply} \\ \text{Effects on Unskilled} \\ \leq 0}} + \underbrace{\left(\Delta L_H^R - \Delta L_H^P\right) r(\tilde{L})}_{\substack{\text{Differential Selection} \\ \text{Into Education}}} \tag{11}$$

The first two terms reveal that supply effects tend to reduce inequality. The intuition is simple. Increasing the share of skilled workers puts downward pressure on their wage. Because the high-income group has a greater fraction of skilled workers than the low-income group, this negative effect will be greater for them (differential supply effects on always skilled). Conversely, supply effects benefit unskilled workers. Because unskilled workers are concentrated at the bottom of the distribution, low-income groups will see a greater rise in their earnings as a result (differential supply effects on unskilled). The third term of the equation highlights another important fact: the distributional effects of education also depend on which *type* of unskilled workers benefits most. If all low-skilled workers benefiting from better education originally come from R, in particular, this will increase inequality. In other words, *who* exactly benefits from educational expansion matters significantly for estimating the distributional effects of education. This mechanism will be particularly important for studying the relationship between education and gender inequality.

2.4. Estimation

I now introduce the methodology used to apply this framework to a growth accounting decomposition. The objective is to estimate the contribution of education to the real earnings growth of different groups p, which differ in their relative supply of workers belonging to skill groups s. A useful way to conceptualize this problem empirically is to formulate it as a counterfactual question: what would have been the distribution of income in 2019, had there been no progress in education since 1980? I propose to estimate this counterfactual in five steps, starting from microdata reporting information on the joint distribution of labor income and education (see appendix B.1 for more details).

2.4.1. Downgrade Education Levels

The first step is to reduce educational attainment to match its distribution in 1980: had there been no human capital accumulation, the distribution of skill groups would be $L^{1980} = (L_1^{1980}, ..., L_m^{1980})$ instead of $L^{2019} = (L_1^{2019}, ..., L_m^{2019})$. In practice, I implement this in the microdata by randomly sampling individuals and downgrading their education levels until reaching the counterfactual. I always give priority to workers whose education is closest to the targeted level. For instance, if the share of workers with no schooling rose from 20% in 1980 (counterfactual) to 40% in 2019 (observed), I construct the counterfactual by first reducing the education of primary-educated workers from primary education to no schooling. I then reduce the education levels of secondary- and tertiary-educated workers only if necessary, until reaching the targeted share of 20%. The resulting database thus contains "untreated" workers, whose education is unchanged, and "treated" workers whose educational attainment is downgraded by one or several levels. This method is similar to the one recently used by Hershbein, Kearney, and Pardue (2020) to simulate the economic effects of expanding access to higher education in the United States.

2.4.2. Reduce Wages Using Returns to Schooling

The second step is to reduce the earnings of treated workers by an estimate of the return to schooling. This implies calculating the return to schooling that should be used, which, as discussed in section 2.2, lies in-between initial and final returns.

Estimation of Initial and Final Returns In the present context, the final return is the return prevailing in 2019, that is, after educational expansion. This return can be estimated in the data using, for instance, a standard Mincerian wage regression.

In contrast, the initial return corresponds to the return that would prevail in 2019, had there been no educational expansion since 1980. This return is unobserved and has to be recovered from the model. With a CES production function:

$$r(L^{1980}) = r(L^{2019}) + \frac{1}{\sigma}\Delta \log\left(\frac{L_H}{L_L}\right) \tag{12}$$

Changes in the relative supply of skilled and unskilled workers are observed, so initial returns can be calculated for any chosen elasticity using this formula. The initial return is higher than the return observed in 2019: if education was to come back to its 1980 levels, the gains from human capital accumulation would appear substantially higher than those observed today.[4]

Estimation of the True Return The true return that should be used lies in-between initial and final returns, and can be calculated using equation 10. Wages and relative supplies are observed, so the only missing parameter to do so is A_H/A_L. Rearranging equation 6:

$$\frac{A_H}{A_L} = \exp\left[\frac{\sigma}{\sigma-1}\log\left(\frac{w_H}{w_L}\right) + \frac{1}{\sigma-1}\log\left(\frac{L_H}{L_L}\right)\right] \tag{13}$$

For a given value of the elasticity, A_H/A_L can thus be recovered from 2019 returns to schooling $\log(w_H/w_L)$ and 2019 relative supplies $\log(L_H/L_L)$ (as in, e.g., Rossi, 2022). This completely closes the model, and allows for a direct calculation of the true return r^* that should be used.

Application to the Data Once the individual return to schooling is estimated, it suffices to directly reduce the earnings of treated workers by the return to schooling. For instance, workers downgraded from primary education to no schooling see their earnings divided by the value of the return to primary education.

2.4.3. Adjust Relative Wages to Account for Supply Effects

The third step is to account for the distributional incidence of supply effects. Changes in relative wages are given by:

$$\Delta\log\left(\frac{w_H}{w_L}\right) = -\frac{1}{\sigma}\Delta\log\left(\frac{L_H}{L_L}\right) \tag{14}$$

Assuming that σ is known, one can directly adjust relative wages in the microdata. For instance, for an increase in the relative supply of skilled workers of 1 log point and an elasticity of substitution of 4, the average wage gap between skilled and unskilled workers is reduced by 0.25 log points. Notice that the aggregate effect of educational expansion is entirely captured in step 2, so the average wage is left unchanged in this step of the estimation.

[4]It is important to stress that this exercise does not amount to calculating the returns that actually prevailed in 1980. As evident from equation 6, returns to schooling are a function of both relative supplies L_H/L_L and the skill bias of technology A_H/A_L. The return observed in 1980 is thus the product of both 1980 relative supplies and 1980 technology. In contrast, the return we are interested in here is the return that would prevail in 2019 if relative supplies were to come back to their 1980 levels, but technology was to remain unchanged at its 2019 level. This return is by construction never observed and has to estimated. I come back to this in section 6.

2.4.4. Derivation of Total Income

Steps 1 to 3 yield a counterfactual distribution of labor income absent educational expansion from 1980 to 2019. The fourth step is to move from this counterfactual distribution of labor income to a counterfactual distribution of total income. Assuming that we know the joint distribution of labor and capital income, this simply amounts to calculating $\tilde{Y}^p = Y_K^p + \tilde{Y}_L^p$ for any given social group or quantile of the income distribution.

2.4.5. Growth Accounting

The final step is to calculate the share of growth in the real income of group p that can be attributed to human capital. This is equal to the gap between the counterfactual and actual growth rate, expressed as a fraction of actual growth from 1980 to 2019:

$$\text{Success}^p = \frac{g^p - \tilde{g}^p}{g^p} \tag{15}$$

With $g_t^p = \frac{y_{2019}^p - y_{1980}^p}{y_{1980}^p}$ and $\tilde{g}_{it} = \frac{\tilde{y}_{2019}^p - y_{1980}^p}{y_{1980}^p}$. If real income growth had been exactly the same absent human capital accumulation, then $\tilde{g}_{it} = g_{it}$ and hence $\text{Success}^p = 0\%$. On the contrary, if there would have not been any growth at all in the absence of educational progress, then $\tilde{g}_{it} = 0$ and $\text{Success}^p = 100\%$: human capital can explain all of economic growth. Notice that as in standard growth accounting, success can be higher than 100%, if actual growth rates fall below those predicted by changes in educational attainment.

3. Data, Methodology, and Stylized Facts

This section presents the data sources and methodology used to estimate the contribution of education to global income and gender inequality reduction since 1980. First, I combine a new set of surveys covering education and wages in 150 countries (section 3.1) with data on the evolution of educational attainment since 1980 (section 3.2). Using estimates of returns to schooling (section 3.3) and of supply effects induced by educational expansion (section 3.4), I estimate by how much lower would wages be, within and between countries, had there been no educational progress since 1980. Finally, I exploit data on the actual evolution of the global income distribution to construct the distributional growth accounting decomposition (section 3.5). Section 3.6 validates the overall methodology by comparing estimates derived from the model to those obtained from the study of three large-scale education policies in India, Indonesia, and the United States.

3.1. Survey Microdata

The starting point is a unique set of household surveys covering the joint distribution of personal income and education by age and gender in 150 countries, which I have assembled for this paper. These surveys come from two main sources.

The first data source is the International Labor Organization's database of labor force surveys. Based on a considerable data collection effort and with the collaboration of national statistical institutes, ILOSTAT have harmonized over 1,300 household surveys, covering 130 countries over the 1990-2022 period. The database records individual-level information on wages, self-employment income, education, and other sociodemographic variables such as age, gender, occupation, industry, and hours worked. All surveys are nationally representative. Most surveys are labor force surveys, which aim specifically at capturing the dynamics of employment and earned income. In some countries where labor force surveys do not exist, the ILO has instead collected consumption or living standards surveys, which generally aim to record household expenditure but also contain information on individual wages and self-employment. In the main analysis, I use the latest survey available in each country.

The coverage of ILO microdata is remarkable, but the information collected on education is limited in some countries. Furthermore, a number of countries are missing, including big countries such as China and Russia. To expand the coverage and quality of the database, I turn to the websites of national statistical institutes and other sources, from which I collect additional household surveys for 55 countries. These include the European Union statistics on income and living conditions (EU-SILC), providing individual microdata for 32 European countries, as well as the Life in Transition Survey (LITS), which covers 10 additional countries in Eastern Europe and Central Asia. I complement these two cross-national datasets with surveys available from country-specific data portals. These sources allow me to cover 13 additional countries: China, Iraq, India, Japan, Mozambique, Morocco, Russia, Somalia, South Africa, South Korea, South Sudan, Tunisia, and the United States. In each case, I collect detailed information on personal income, education, and other sociodemographic variables, which I harmonize in the same way as the ILO. More details can be found in appendix E.

Figure 2 maps the geographic coverage of the resulting database. Table 1 provides descriptive statistics. The data cover about 9.6 million individuals surveyed in 150 countries. These surveys are representative of over 95% of the population of each world region. The exception is the Middle East and North Africa, where microdata is well known to be either non-existent or inaccessible to researchers (Ekhator-Mobayode and Hoogeveen, 2022). Overall, the microdata cover about 95% of the world's population, and about 93% of the world's GDP.[5] To the best of my knowledge, this represents the most comprehensive micro-database on the world distribution of income ever built in economics research.

3.2. Educational Attainment Data

The first step of the estimation consists in downgrading the education of individuals to match its distribution observed in 1980. This requires data on the evolution of educational attainment in each country. Here, the primary source is the database compiled by Barro and Lee (2013) and updates.[6] It records estimates of the share of individuals with no schooling, primary education, secondary education, and higher education by 10-year age groups and gender, in 146 countries, from 1950 to 2015. A number of countries are covered by the survey microdata but absent from the Barro-Lee database. For these, I construct my own estimates, using either census data from IPUMS International, cohort-level trends observed in the labor force surveys, or other sources (see appendix F).

For the counterfactual to be valid, it is important to ensure that education categories recorded in surveys match those observed in the Barro-Lee database. This is particularly challenging for a number of reasons. For instance, the ILO sometimes includes incomplete degrees in each education level and sometimes does not, depending on the information available in each survey. Meanwhile, the Barro-Lee database sometimes includes lower secondary education with primary education. Comparing the distribution of educational attainment in the two sources, I manually map categories in each country, one by one, until estimates from the two datasets coincide as close as possible. Finally, to ensure that the construction of the counterfactual is perfectly consistent, I re-calibrate the sample weights in each survey to make them match the distribution of educational attainment by age and gender recorded in the Barro-Lee database. This last step only marginally affects weights, given that the distribution of educational attainment is very close in the two sources after reclassification.[7]

Figure 3 plots the distribution of global educational attainment from 1980 to 2019. There has been a dramatic expansion of secondary education, from about 20% to nearly 60%. This

[5] In nearly all countries, the survey was fielded after 2015, ensuring that the database is broadly representative of the global distribution of income over the 2015-2019 period. Appendix figure E1 provides a map of the corresponding survey years in each country.

[6] See `https://barrolee.github.io/BarroLeeDataSet/BLv3.html`.

[7] Appendix figures F1, F2, F3, and F4 compare estimates of the share of the working-age population with no schooling, primary education, secondary education, and tertiary education in the Barro-Lee database and the survey microdata, after manually mapping educational categories in the two sources.

rise was mirrored by a significant decline in the share of adults with primary education or no schooling. In 2019, less than 15% of working-age adults had not attended at least some years of basic education. Tertiary education also expanded significantly, from less than 5% in 1980 to over 10% in 2019.[8]

3.3. Returns to Schooling

To estimate how much lower incomes would be absent educational progress, one needs estimates of returns to schooling in 2019 (corresponding, as discussed in section 2.4, to final returns). There are two options, each of which with advantages and disadvantages: returns to schooling estimated by OLS in my data, or causally identified returns available from the existing literature.

3.3.1. OLS Returns

In the main analysis, I rely on OLS estimates of returns to schooling by education level, derived from a modified Mincerian equation of the form:

$$\ln y_{ic} = \alpha_c + \beta_c^{pri} D_{ic}^{pri} + \beta_c^{sec} D_{ic}^{sec} + \beta_c^{ter} D_{ic}^{ter} + X_{ic}\beta_c + \varepsilon_{ic} \tag{16}$$

With y_{ic} earned income of individual i in country c, D_{ic}^{pri}, D_{ic}^{sec}, and D_{ic}^{ter} dummies for having reached primary, secondary, and tertiary education, and X_{ic} a vector of controls including gender, an experience quartic, and interactions between gender and the experience quartic (as in Lemieux, 2006; Autor, Goldin, and Katz, 2020). I restrict the sample to all individuals with positive personal income, including both wage earners and self-employed individuals. The dependent variable is total annual earned income from all jobs. This ensures that the estimated returns to schooling are not sensitive to effects of education on hours worked, which have been shown to be significant and variable across countries (Bick, Fuchs-Schündeln, and Lagakos, 2018). I assume that education does not affect aggregate employment. This is probably a conservative assumption, given existing evidence on positive effects of schooling on female labor force participation (see section 5). Appendix G provides results for other empirical specifications, variable definitions, and sample restrictions; the resulting returns are largely insensitive to alternative methodological choices.

The main advantage of OLS returns lies in their coverage and comparability: they can be

[8] Improvements in schooling have coincided with significant declines in overall educational attainment inequalities, both between and within countries: see appendix figure A20, which provides a Theil decomposition of global educational attainment inequality following the methodology first proposed by Morrisson and Murtin (2013). Appendix figure A21 charts average years of schooling of the working-age population by world region since 1980.

estimated for all 150 countries using a unified methodology. The potential disadvantages are twofold.

A first source of concern is that returns estimated by OLS may suffer from omitted variable bias. A classical argument is that ability bias may lead to overestimating the return to schooling: more educated workers have higher earnings because of greater innate abilities, not because of schooling itself. Reassuringly, several decades of labor economics literature teaches us that omitted variable bias is typically small and, if anything, tends to go in the opposite direction (for reviews, see Card, 1999; Deming, 2022; Gunderson and Oreopolous, 2020). Returns estimated by OLS are thus likely to provide a good approximation of average individual returns to schooling, if not a lower bound.

A second source of concern is that our parameter of interest is not the *average* return to schooling: it is the return to schooling for those who benefited from educational expansion since 1980. Consider a country in which the share of workers with primary education increased from 10% to 20% between 1980 and 2019. Intuitively, this implies that by 2019, 10% of workers correspond to workers who would have obtained primary education anyway in a world with no educational expansion (the "always skilled"), while another 10% benefited from increased access to schooling (the "newly skilled"). One would like to estimate the return for these 10% of newly skilled, yet an OLS estimate over the entire sample will capture both groups. Given potentially large heterogeneity in returns across individuals (e.g., Heckman, Humphries, and Veramendi, 2018), our parameter of interest may differ significantly from the one estimated over the entire population. In particular, improved access to schooling has predominantly benefited children from lower socioeconomic backgrounds (Gethin, 2023), a population that is often found to have higher returns. The next section provides evidence that this is indeed likely to be the case.

Figure 4 plots the resulting estimates of returns to schooling by world region. Returns to schooling are strongly convex, in particular in low-income countries. In Sub-Saharan Africa, the average return to a year of primary education is 6%, while the average return to a year of tertiary education is 22%. Returns to primary education are extremely low in India, barely reaching 3%. A direct implication is that using the average return would likely lead to overestimating the contribution of education to global poverty reduction, given that the global poor have mostly benefited from expansions in access to basic education, which displays the lowest returns. There are also significant variations in average returns across regions. Europe and the United States have the highest average return to schooling, at over 12%, while it falls below 6% in the Middle

East and North Africa.[9]

3.3.2. IV Returns

An alternative option is to use instrumental variable estimates of returns to schooling, in particular those derived from differential exposure to large-scale education programs. The estimation of these returns generally relies on comparing cohorts or regions that were more or less exposed to specific policies, such as compulsory schooling laws or school construction programs. They are causally identified, so they do not suffer from omitted variable bias. Another major advantage is that they focus by construction on the newly skilled, since they are based on comparing the earnings of those who marginally gained access to education to those who did not. The main drawback is that they can only cover specific programs expanding access to specific levels of education (although the estimates compiled below do end up covering about 60% of the world's population).

The labor economics literature has made considerable efforts at expanding such estimates to multiple contexts and policies in recent years, in particular in developing countries. For the purpose of this paper, I have assembled a collection of IV estimates of the returns to schooling, drawing on a number of recent studies.

I select estimates based on four criteria. First, I give priority to articles studying episodes of large-scale expansions in access to schooling. Second, I select studies for which a comparable OLS estimate is available for comparison. Third, I restrict the sample to relatively recent studies, covering policies that expanded access to education during my period of interest (see Card (1999) for similar findings based on older studies mostly conducted in rich countries). Fourth, I select one estimate per country and education level.

Figure 5 plots the resulting OLS and IV estimates of the return to schooling. Two results stand out. First, OLS and IV returns are highly correlated. This provides reassuring evidence that OLS estimates capture true variations in returns to schooling across countries and education levels relatively well. Second, IV returns are almost always higher than OLS returns or not significantly different (in line with previous estimates compiled by Card (1999) for developed countries). In some countries, such as China, IV returns are two to three times larger. Appendix figure G7 plots the ratio of IV to OLS estimates across studies. The gap ranges from close to 0%

[9] Appendix figure G2 plots the cross-country distribution of returns to schooling, while appendix figures G3, G4, G5, and G6 map average returns to schooling and returns to primary, secondary, and tertiary education in all countries with available data. The median returns to primary, secondary, and tertiary education are 5%, 9%, and 13%, respectively. See also figure G1, which presents results of cross-country regressions by education level and further decomposes secondary education into upper secondary and lower secondary.

in Nigeria to almost 250% in China (tertiary). The average gap is in the order of 80%.

As mentioned above, I use OLS returns in the main analysis. As an alternative specification, I exploit IV estimates to correct returns to schooling upwards in three steps. First, I multiply OLS estimates by the ratios plotted in appendix figure G7 for each country-level covered by the data (for instance, I increase the return to tertiary education in Vietnam by 50%). Second, I multiply the return to other levels of education by the same ratio (for instance, I increase the returns to primary and secondary education in Vietnam by 50% too). This amounts to assuming that returns to other levels of schooling are underestimated by the same factor in each country. Third, I increase returns to schooling in missing countries by the average correction factor observed, that is, 80%.

3.4. Supply Effects

The third step of the methodology is to account for supply effects, which are necessary to estimate the true return to schooling (section 2.2) and the distributional effects of schooling expansion (section 2.3). This implies extending the CES production function to more than two levels of schooling, and calibrating elasticities of substitution between skill groups.

3.4.1. Production Function Specification

Until now, I have worked with a CES production function with two skill groups, yet the data provide the distribution of educational attainment for four: workers with no schooling, primary education, secondary education, and tertiary education. To incorporate supply effects on these four skill groups, I introduce nests in the CES production function, in line with previous work in labor economics (see in particular Goldin and Katz (2007) on the United States and Fernández and Messina (2018) on Latin America).

At the top level, output is produced by combining workers with tertiary education and workers with below tertiary education:

$$L = \left(A_{\bar{ter}} L_{\bar{ter}}^{\frac{\sigma_1 - 1}{\sigma_1}} + A_{ter} L_{ter}^{\frac{\sigma_1 - 1}{\sigma_1}} \right)^{\frac{\sigma_1}{\sigma_1 - 1}} \tag{17}$$

With L_{ter} the share of college-educated workers and $L_{\bar{ter}} = 1 - L_{ter}$. The subgroup of workers with less than tertiary education is split into workers with secondary education and workers with below secondary education:

$$L_{\bar{ter}} = \left(A_{\bar{sec}} L_{\bar{sec}}^{\frac{\sigma_2 - 1}{\sigma_2}} + A_{sec} L_{sec}^{\frac{\sigma_2 - 1}{\sigma_2}} \right)^{\frac{\sigma_2}{\sigma_2 - 1}} \tag{18}$$

With L_{sec} the share of secondary-educated workers and $L_{\bar{sec}} = 1 - L_{sec} - L_{ter}$ the share of workers with less than secondary education. This corresponds exactly to the specification adopted by Goldin and Katz (2007) and Fernández and Messina (2018). Finally, because my data also include countries with a significant fraction of workers with below primary education, I introduce a third nest separating workers with primary education from those with no schooling:

$$L_{\bar{sec}} = \left(A_{non} L_{non}^{\frac{\sigma_3 - 1}{\sigma_3}} + A_{pri} L_{pri}^{\frac{\sigma_3 - 1}{\sigma_3}} \right)^{\frac{\sigma_3}{\sigma_3 - 1}} \tag{19}$$

With L_{pri} the share of workers with primary education and L_{non} the share of workers with no schooling. This cutoff has been adopted to study episodes of expansions in access to primary education (see in particular Khanna (2023) on India).

These three equations yield three formulas for the returns to primary education, secondary education, and tertiary education, which map directly onto the categories reported in the Barro-Lee database and the returns to schooling plotted in figure 4. True returns to schooling (equation 10) and relative wage adjustments (equation 14) can then be calculated separately for each of the three nests.[10]

3.4.2. Elasticities of Substitution

To close the model, the only parameters that need to be calibrated are the elasticities of substitution between skill groups. A rich empirical literature in labor economics has attempted to estimate these elasticities in various contexts, with resulting values ranging from 1.5 to 5 and typically close to 2, depending on the methodology used, the chosen skill cutoff, and the country considered. Appendix table B1 reports selected estimates from existing empirical studies.

Most studies rely on short-run variations in relative skill supply to estimate elasticities of substitution. As a result, they identify a relatively short-run elasticity, corresponding to the case in which the skill bias of technology (A_H/A_L) is held approximately constant. Yet in the long run,

[10] Appendix figures B6, B7, and B8 plot relative skill efficiency for the three levels of the production function versus GDP per capita. In line with recent evidence by Rossi (2022), there is a strong correlation between the relative efficiency of skilled labor and economic development: skilled workers are substantially more efficient in rich countries. The United States stand out as displaying a particularly large relative efficiency of tertiary-educated workers. This directly results from the return to tertiary education being very large in spite of the U.S. having a high share of tertiary-educated workers.

one should expect firms to adjust their technological mix as a response to the greater supply of skilled workers (e.g., Acemoglu, 1998). This would imply less sensitivity of relative wages to relative supplies, and hence greater values of the elasticity of substitution after accounting for endogenous technical change.

Motivated by this fact, a recent literature in macroeconomics attempts to estimate long-run elasticities of substitution between skill groups. Two recent articles, in particular, have made significant progress in this direction. Exploiting data on wage gains at migration to the United States for different skill groups, Hendricks and Schoellman (2023) find a long-run elasticity ranging from 4.5 to 8, depending on the cutoff chosen to define high-skilled workers. Bils, Kaymak, and Wu (2022) instead exploit data on worldwide trends in education and returns to schooling to pin down values of the elasticity of substitution consistent with both no worldwide technological regress and a greater skill bias of technology in rich countries. These two conditions yield lower and upper bounds of 4 and 6, very similar to the numbers obtained by Hendricks and Schoellman (2023).

To the extent that this article analyses large improvements in educational attainment over several decades, long-run elasticities of substitution are the relevant parameters. In my main analysis, I thus calibrate $\sigma_1 = \sigma_2 = \sigma_3 = 6$, corresponding to the midpoint of estimates reported in Hendricks and Schoellman (2023) and Bils, Kaymak, and Wu (2022). I discuss the robustness of my results to alternative specifications in section 4.4.

3.5. Global Labor and Capital Income Inequality Data

The final step of the estimation consists in moving from labor income to total income, and comparing counterfactual to actual real income growth rates. This requires data on the distribution of income in each country, aggregate labor and capital income shares, and the share of income received from labor and capital by income group within each country.

3.5.1. Global Income Inequality Data

Data on the world distribution of income come from the World Inequality Database (WID). The database covers average per-capita income by percentile in all countries in the world, every year from 1980 to 2019. The income concept is pretax national income, that is, total income received by individuals before accounting for taxes and transfers, but after accounting for the operation of pension and unemployment systems. Importantly, all components of net national income (GDP, minus consumption of fixed capital, plus net foreign income) are allocated to individuals, following the Distributional National Accounts (DINA) framework (see Chancel et al., 2022; Piketty, Saez, and Zucman, 2018). This ensures that all income distributions are consistent with macroeconomic growth rates and aggregate capital and labor income shares recorded in the national accounts. The database is constructed by compiling estimates from detailed national or regional studies, which combine surveys, tax data, and national accounts to construct distributions that are conceptually comparable across countries (see for instance Piketty, Saez, and Zucman (2018) on the United States, Blanchet, Chancel, and Gethin (2022) on Europe, and De Rosa, Flores, and Morgan (2022) on Latin America).

3.5.2. Aggregate Labor and Capital Income Shares

Aggregate factor income shares come from Bachas et al. (2022), who combine a number of sources to build a new database on the components of net national income worldwide since 1965. Their database provides a decomposition of net domestic product into compensation of employees, mixed income, the operating surplus of households (actual and imputed rental income), and the operating surplus of corporations (profits net of depreciation).

I define the labor income share as the share of income attributable to compensation of employees and mixed income. This is the definition of the labor share that is the most conceptually meaningful in my context, given that my microdata cover individual income and returns to schooling for both wage earners and the self-employed. In the main analysis, I thus make the conservative assumption that human capital only affects wages and mixed income, while leaving capital income unchanged.

3.5.3. Capital Income Concentration

The last step is to estimate how the capital income share varies alongside the income distribution in each country. High-quality data on this decomposition are scarce, given well-known issues with the underestimation of capital income in household surveys. Drawing on the few studies that were able to mobilize tax and national accounts data to estimate such decomposition with a relatively good level of precision, I was able to derive profiles of the capital income share by percentile for the United States (Piketty, Saez, and Zucman, 2018), South Africa (Chatterjee, Czajka, and Gethin, 2022), and 10 Latin American countries (De Rosa, Flores, and Morgan, 2022). The corresponding series are plotted in appendix figure A19. The profiles look very similar across these three cases. The capital share is always below 20% for the bottom 90% of earners, corresponding mostly to imputed rental income. It rises exponentially at the very top of the distribution, where the main source of income is from bonds and stock. Given these similarities, I use the average profile observed across countries, which I rescale in each country-year to match the aggregate capital income share.

3.6. Validation from Three Natural Experiments

Despite the relative popularity that the canonical labor demand and supply framework has encountered in the literature, the ability of my estimates to capture the true aggregate and distributional effects of educational expansion may naturally be questioned. To shed light on the validity of this approach, I turn to causal evidence from three natural experiments. I focus on outlining the main results. The interested reader will find more details in appendix C.

3.6.1. Contexts, Data, and Methodology

I investigate the distributional effects of three large-scale education policies: India's District Primary Education Program (1990s-2000s), Indonesia's INPRES school construction program (1970s), and U.S. state compulsory schooling laws (1870s-1960s). These three sets of policies have been extensively studied to estimate individual returns to schooling, human capital externalities, and general equilibrium effects affecting different skill groups (e.g., Acemoglu and Angrist, 2000; Duflo, 2001; 2004; Khanna, 2023). Less is known of their exact distributional effects by income group.

Combining data from pre-existing studies and additional sources, I first exploit these natural experiments to estimate the causal effect of educational expansion on the distribution of income at the level of subnational regions. More specifically, I run variants of the following specification:

$$\ln y_{rt}^i = \gamma_0^i + \gamma_1^i S_{rt} + X_{rt}^i \beta + \delta_r + \delta_t + \varepsilon_{rt} \tag{20}$$

$$S_{rt} = \alpha_0 + \alpha_1 Z_{rt} + \eta_{rt} \tag{21}$$

Where y_{rt}^i denotes the average income of income group i (such as the bottom 20% of earners) in subnational region r at time t. The objective is to estimate the impact of an exogenous increase in S_{rt}, the average years of schooling of the working-age population in region r. X_{rt}^i is a vector of controls, such as the demographic composition of the region, δ_r are subnational region fixed effects, and δ_t are time fixed effects. The parameters of interest are γ_1^i for different groups i, which provide reduced-form estimates of the effect of increasing average years of schooling on the distribution of income.

S_{rt} is instrumented by Z_{rt}, a variable capturing quasi-experimental variation in exposure to the education program. In India, I rely on Khanna (2023), who estimates the impact of the DPP using a regression discontinuity design around the cutoff district literacy rate used to allocate the program. In Indonesia, I instrument district average years of schooling by the number of schools built under the INPRES program, following Duflo (2001). In the United States, the instrument is average required years of schooling across cohorts born in different states, in the spirit of the existing literature (e.g., Guo, Roys, and Seshadri, 2018).

After estimating the effects of each program, I then compare these results to aggregate and distributional effects predicted by the model. In India, for instance, I simulate the effect of increasing average years of schooling by one year through primary education, following each step of the methodology outlined in section 2.4. This yields simulated estimates of γ_1^i, which can be compared to those obtained empirically from the natural experiment.

3.6.2. Main Results

Figures 6, 7, and 8 plot the main results, comparing the estimated and simulated effects of educational expansion on the average income of each income quintile. All three policies led to large reductions in income inequality. In India, for instance, increasing average years of schooling by one year in a treated district is associated with a 20% increase in the average income of the bottom quintile, compared to a null effect on the average income of the top 20%. Aggregate effects of education on earnings are found to range from 8% to 15% per average year of schooling (as shown by the dashed line in each figure), which is relatively close to individual Mincerian returns.

The model performs remarkably well at reproducing results from these natural experiments. In all three cases, simulations predict significantly higher returns to schooling expansion at the bottom of the distribution. This is because these policies targeted basic education, which disproportionately benefits low-income earners in the simulation, and because supply effects magnify this redistributive effect. If anything, the model slightly underestimates benefits at the bottom of the distribution, in particular in the case of the United States. Together, these results provide reassuring evidence that the methodology developed in this paper performs well at estimating the distributional effects of educational expansion, and may even provide a lower bound on true benefits for low-income earners.

4. Education and the World Distribution of Income

This section presents the main results on the role of education in reducing global poverty and inequality. Section 4.1 focuses on the overall distribution of global economic growth since 1980.

Section 4.2 decomposes the effects of education by world region, time period, and between and within countries. Section 4.3 compares the results to those of a standard growth accounting decomposition, isolating the contribution of each step of the distributional growth accounting methodology. Section 4.4 provides various robustness checks and extensions.

4.1. Education and the Distribution of Global Economic Growth

I start by presenting results on the role of education in shaping the distribution of global economic growth since 1980. Table 2 presents a distributional growth accounting decomposition of the world distribution of income for the 1980-2019 period.

1) Education Explains 50-60% of Average Economic Growth Global average income per capita approximately doubled over this period (+98%). Absent educational expansion, growth would have been significantly lower, at 45%. Education thus contributed 53 percentage points of growth. Taking the ratio between the contribution of education and actual economic growth, private returns to schooling explain about 54% of average per capita income growth since 1980.

2) Education Explains 60-70% of Growth for the Global Poorest 20% This average figure hides significant heterogeneity by global income group. For the poorest 50% of individuals in the world, growth has been markedly higher, exceeding 150%. Yet, real income gains from educational expansion have also been higher for this group, so that the share of growth explained by education reaches almost 60%. Overall, private returns to schooling can account for 60% to 80% of growth for the global bottom 90%. This share is highest for the bottom 20% (69%) and middle 40% of the income distribution (78%), two groups that have witnessed lower growth and large gains from increased access to schooling. It is lowest at the very top of the distribution, mainly because the bulk of income at the top is received from capital (which by assumption is not affected by schooling).

Figure 1 provides a more granular picture of the distribution of global economic growth since 1980. All individuals in the world are ranked from the poorest 1% to the richest 0.01%. Total pretax income growth is then calculated for each percentile, together with growth explained by private returns to schooling (lower shaded area) and residual growth coming from other factors (upper shaded area). Real income gains have been greatest at the middle and the very top of the global income distribution, generating what has often been referred to as the "elephant curve" of global inequality and growth (Lakner and Milanovic, 2016). This pattern reflects the conjunction of trends in inequality between and within countries, including the rise of China and

India (middle of the distribution), sluggish economic growth in low-income countries (bottom of the distribution), weak income gains for most households living in high-income countries (upper middle of the distribution), and skyrocketing top income inequality in many parts of the world (top end of the distribution).[11] The main contribution of this paper is to isolate gains from education, represented by the lower shaded area. These gains have been particularly large for most income groups, ranging from 80 points to 120 points for most percentiles within the bottom 90%.

Taking the ratio of the contribution of education to actual growth rates yields figure 9, which represents the share of growth explained by education by global income percentile. This share ranges from 55% to 95% for all income groups within the bottom 90%. It is highest at the bottom and the upper middle of the income distribution, exceeding two-thirds for the poorest 20% and for groups ranging from the 70th to the 90th percentiles.

3) Education Explains 40-70% of Global Poverty Reduction Beyond growth for specific groups, another indicator that has received considerable attention is the share of the world's individuals living in extreme poverty. A difficulty in the context of this paper is that poverty headcount ratios are based on counting the number of individuals whose income falls below a certain threshold rather than on actual income gains. This makes the calculation of the share of poverty reduction explained by education less conceptually meaningful (since it implies counting people rather than comparing growth rates) and more sensitive to the choice of a specific threshold (given that poverty rates are not necessarily linear in growth rates).

Another limitation is that estimated poverty rates can differ significantly across sources. Most commonly used estimates are those provided by the World Bank, but these are not ideal in the context of this paper for two main reasons. First, they rely exclusively on data from household surveys, which generally miss capital income entirely (Blanchet, Flores, and Morgan, 2022) and can display growth rates that differ substantially from those reported in the national accounts (Pinkovskiy and Sala-i-Martin, 2016). Second, they are based on household expenditure rather than pretax income, which may bias the results depending on the size and distributional incidence of taxes, government transfers, and saving rates. An alternative solution is to rely on pretax income distributions from the World Inequality database, which have the advantage of covering the correct income concept and being consistent with macroeconomic growth rates.

[11]The interested reader will find additional figures describing these trends in more detail in the appendix. In particular, appendix figure A13 provides a Theil decomposition of global income inequality since 1980, while figure A14 charts the average share of pretax income received by the richest 10% by world region. Figures A15, A16, A17, and A18 provide further graphical evidence on the geographical breakdown of global income groups in 1980 and 2019, as well as the geographical composition of the world's poorest 20% and richest 20% since 1980.

The difficulty is that poverty thresholds were designed by the World Bank to match deprivation levels reported in surveys, not GDP per capita levels.

With these limitations in mind, table 3 extends the growth accounting decomposition to global poverty headcount ratios at $2.15, $3.65, and $6.85 per day (the three thresholds typically used by the World Bank), calculated using pretax income distributions from the World Inequality Database. Poverty at $2.15 per day fell from 20% in 1980 to 9% in 2019. Absent educational expansion, it would be about 4 percentage points higher today, implying a decline in the global poverty headcount ratio of 32% instead of the 55% observed. By this measure, private returns to schooling can account for 43% of global extreme poverty reduction. The corresponding figures are 28% at $3.65 per day and 43% at $6.85 per day.

As an alternative, I reproduce this analysis using World Bank poverty rates.[12] The results are presented in appendix table A3. Education explains 39% of global poverty reduction at $2.15 per day, 58% at $3.65 per day, and 75% at $6.85 per day, similar to or higher than the figures obtained with the WID data. The takeaway is again that education has been a major driver of improved living standards for the world's poorest individuals.

4.2. Decomposing Global Schooling Gains

Faced with these results, one may naturally wonder what are the different factors driving them. Why does education explains a higher share of growth at the bottom and upper-middle of the world distribution of income? Are the results primarily driven by the distribution of growth within countries or by differences in aggregate gains from schooling across countries? This section attempts to answer these questions.

4.2.1. Distributional Growth Accounting by World Region

A first way of better understanding the results is to decompose them by world region. Table 4 displays growth decompositions for selected geographical regions and country income groups, distinguishing between average economic growth and real income growth of the poorest 20%.

[12] The World Bank does not publish data on the world distribution of income. I thus reconstruct it myself by collecting income and consumption distributions from the World Bank's website and extrapolating the average income of each country-percentile to missing years using real GDP per capita growth rates. This yields trends in global poverty almost identical to those officially reported by the World Bank. Finally, I construct counterfactual income distributions using the same methodology as in the rest of the paper. The main difference is that capital income is absent from World Bank surveys, implying a 100% passthrough of schooling on income instead of a passthrough equal to the labor share. In addition to the poverty analysis, appendix table A2 reproduces the main distributional growth accounting decomposition using World Bank data. With this dataset, education accounts for about 60% of global bottom 20% growth since 1980.

Two main results stand out.

1) Education Explains 60-100% of Growth in Low- and High-Income Countries Looking at aggregate growth figures, education explains the totality of growth for low-income countries and can explain about 59% of growth for high-income countries. Sub-Saharan Africa and Latin America are the two world regions where education explains the highest share of growth (over 100% and 82% respectively), despite the fact that gains from education have been the lowest. The reason is simply that these are the two regions that have witnessed the lowest average growth rates over the period considered. In contrast, private returns to schooling alone cannot account for the exceptional growth rates witnessed by China and India, despite substantial improvements in schooling in these two countries. Interestingly, although economic growth has been about two times slower in India than in China, education can only explain 26% of it in India compared to 32% in China. The main reason is that a significant fraction of educational expansion in India has occurred at the level of basic education, which displays exceptionally low returns (figure 4), while China has primarily benefited from expansions in secondary and tertiary education.

2) Education Explains Over 40% of Growth For Low-Income Groups in All Regions In each region or country, education almost always explains more growth at the bottom of the distribution than for the average individual, for two main reasons. First, because of rising inequality in many countries, actual growth has been significantly lower at the bottom of the distribution (in particular in Europe, Northern America, China, and India). Education thus explains more of growth at the bottom of the distribution, simply because there is less growth to be explained. Secondly, gains from schooling have been greater for low-income earners than for high-income earners in most world regions, mainly because supply effects tend to reduce inequality and because capital income is concentrated at the top of the distribution (I come back to this in section 4.3). As a result, education can explain more than 100% of growth for the bottom 20% of earners in Western economies, Latin America, MENA, and Sub-Saharan Africa. Even in India, which displays extraordinarily low returns to basic education, private returns to schooling can explain 44% of real income gains for the poorest 20%.

Combining these two facts enables a better understanding of the patterns presented in figure 1. Economic growth has been the highest for Chinese and Indian middle classes, corresponding to income groups near the median of the global income distribution. Although these groups have benefited from some of the highest gains from schooling, these gains cannot fully account for such exceptional growth rates. At the upper-middle of the global income distribution, stagnating real incomes for European and US low-income earners have coincided with relatively high gains from schooling for these groups, which is why education explains the bulk of growth for percentiles 70 to 90. Finally, growth has been weak for the world's poorest individuals, mainly due to low economic growth in Sub-Saharan Africa, but also because of rapidly rising inequality in middle-income countries (in particular India). Schooling gains have not been particularly impressive in low-income countries either, but they have been high for low-income earners of both low- and middle-income countries. This explains why education can account for such a large share of growth at the bottom of the global income distribution.

4.2.2. Distributional Growth Accounting Within and Between Countries

Given these complex patterns of educational expansion affecting inequality within and between countries, has education been a driver of higher or lower global income inequality overall? One way of answering this question is to perform a Theil decomposition of global inequality into a between-country and a within-country component. This decomposition is reported in table 5.

1) Education Has Prevented the Rise of Global Inequality Since 1980 The first two rows of table 5 compare the evolution of the Theil index of worldwide inequality to its counterfactual evolution absent schooling expansion. From 1980 to 2019, global inequality more or less stagnated. Absent educational progress, it would have instead risen dramatically, from 1.06 to 1.35. Educational progress has thus contributed to strongly reducing global inequality in the past decades.

2) Education Has Reduced Inequality Within Countries, But Not Between Countries The next two rows of table 5 compare actual and counterfactual trends in inequality between countries, measured by the between-country component of the Theil index. Education has not had much effect on cross-country income convergence. Inequalities between countries declined strongly, from 0.6 to 0.34, mainly because of the rise of China and India. This decline would have been the same absent educational progress. This result mirrors the complex patterns highlighted above, with education explaining more growth in low-income and high-income countries than in middle-income countries. The effect of education on the overall dispersion of cross-country average incomes appears to have been close to zero.

While the effect of education on between-country inequality is unclear, schooling has been an unambiguous driver of convergence within countries. The Theil index of within-country inequalities grew by 0.28 points over the 1980-2019 period, from 0.46 to 0.74. This increase would have been twice as large absent educational expansion. In other words, education has sufficiently mitigated the rise of within-country inequality to keep overall global inequality constant.[13]

4.2.3. Distributional Growth Accounting by Time Period

The structure of the data also allows me to estimate the contribution of education by time period. This analysis delivers two main results.

1) Schooling Gains Have Increasingly Benefited the Global Poor First, the benefits of worldwide improvements in schooling have increasingly accrued to the global poor since 1980. The best way to see this is to compare the distribution of schooling gains over the 1980-2019 and 2000-2019 periods. Figure 10 plots the contribution of education to the annual income growth of each global percentile for these two periods. The figure is obtained by comparing the distribution of income in 2019 to its counterfactual distribution absent educational progress, and then annualizing the resulting ratio.[14]

Consistently with previous results, schooling gains have been greatest at the middle of the global income distribution since 1980, generating a 1.4% annual increase in earnings for percentiles 60-70, compared to 1.2% for the first quintile and less than 0.8% for all groups within the top decile. In the recent period, growth effects of education have been more clearly progressive, reaching 1.4-1.6% for the first decile, 1-1.2% for percentiles 20 to 70, and less than 0.5% for the top decile. The past decades have thus seen an acceleration of human capital accumulation for the global poor and a relative slowdown of educational progress in the rest of the world. Since 2000, education has been a major driver of global inequality reduction.

2) Education Explains Over 50% of Global Bottom 20% Growth in All Periods Actual economic growth has also accelerated since 2000, a period characterized by the surge of the Chinese and Indian economies and higher growth rates in most other parts of the world. Appendix table A1 presents a distributional growth accounting decomposition of the world distribution of income for the 2000-2019 period. Education can explain 28% of global average economic growth since 2000, about two times lower than the corresponding figure since 1980.

[13] Appendix figure A3 compares average gains from schooling to gains for the poorest 20% individuals in each country. In nearly all countries, education appears to have increased income more for low-income than high-income earners, mainly because of supply effects redistributing income from high-skilled to low-skilled workers.

[14] Formally, let y be income in 2019 and $\tilde{y}$ counterfactual income. The contribution of education is then equal to $(\frac{y}{\tilde{y}})^{(1/T)} - 1$, with T the number of years corresponding to the period considered (40 years for 1980-2019, 20 years for 2000-2019). Notice that my estimates always rely on comparing a given year to 2019, given that I only have microdata for the latter. This implies that I unfortunately cannot estimate the distributional incidence of educational expansion over the 1980-2000 period, for instance.

For the global bottom 20%, however, education can still account for 55% of income gains, because actual growth for this group has not been particularly high, while gains from schooling have increased. The same pattern holds over the 2010-2019 period.[15] All in all, private returns to schooling can always explain more than half of growth for the world's poorest individuals, regardless of the period considered.

4.2.4. Distributional Growth Accounting Across Cohorts

The results presented until now focus on changes in educational attainment of the working-age population from 1980 to 2019. This involves comparing education for cohorts born between the 1910s and 1960s with those born between the 1960s and 1990s. The effect of education on economic growth thus results from two separate mechanisms: the replacement of new cohorts by old ones (e.g., 1910s cohorts are observed in 1980 but not in 2019), and the fact that new cohorts are more educated than their elders (e.g., 1990s cohorts are more educated than 1980s cohorts). In this section, I isolate the specific contribution of the second channel, namely, educational progress among new generations that arrived on the labor market from 1980 to 2019. This approach is interesting from a historical and policy point of view, because it allows answering the following question: how much lower would incomes be if cohorts arriving on the labor market after 1980 had not been more educated than the 1980 cohort? Put simply, it allows capturing the specific contribution of improvements in education among new generations since 1980.

To derive this counterfactual, I use the same methodology as in the rest of the paper, with the only difference that counterfactual educational attainment is that of the 1980 cohort instead of that of the 1980 working-age population. I start by calculating education of the 1980 cohort, estimated as that of individuals aged 60 to 65 in 2019. I then downgrade education levels of each post-1980 cohort until reaching the 1980 counterfactual. Finally, I reduce earnings using returns to schooling and estimate supply effects as in the rest of the paper.

The main results of this exercise are presented in appendix table A7. Appendix figures A4 and A5 present the corresponding distribution of gains from schooling and share of growth explained by global income percentile. Worldwide educational progress appears to have been particularly progressive when narrowing the focus to post-1980 generations. Generational progress explains about 12% of global economic growth since 1980, but 61% of income gains for the world's poorest 20% individuals, 37% for the global bottom 50%, and 4% for the global

[15]Appendix figure A1 plots the share of average growth and global bottom 20% growth that can be explained by education over the 1980-2019, 1990-2019, 2000-2019, and 2010-2019 periods. Appendix figure A2 represents the distribution of global economic growth together with the contribution of education in 2000-2019, as in figure 1.

top 10%. These results are in line with those of the previous section, which highlighted the increasingly progressive nature of educational progress. Given substantial improvements in school enrollment observed in low-income countries in the 1990s and 2000s (Barro and Lee, 2015), this pattern can only be expected to intensify in the future. In the coming decades of the twenty-first century, education could well become an even stronger force of decreasing global inequality that it already was at the turn of the twentieth century.

4.3. Standard Versus Distributional Growth Accounting

Another way of better understanding the results is to compare my estimates to those obtained before and after applying each of the estimation steps outlined in section 2. This is useful to isolate the different mechanisms driving the results, from differential returns to schooling to changes in within-country inequality and general equilibrium effects. It also enables comparing my results to those one would obtain from a canonical growth accounting decomposition.

Table 6 displays the share of global average economic growth and global bottom 20% income growth that can be explained by education with different data sources and assumptions.

4.3.1. The Standard Growth Accounting Decomposition

I start by presenting results from a standard growth accounting decomposition in its simplest form. To do so, I follow the same methodology as Barro and Lee (2015), who exploit cross-country GDP and educational attainment data to estimate the fraction of global economic growth explained by human capital accumulation from 1960 to 2010.[16] This decomposition only requires three ingredients: cross-country per-capita GDP data (taken from the World Inequality Database), capital income shares (taken from the Penn World Tables as in Barro and Lee, 2015), and an estimate of the Mincerian return to schooling (assumed to be 10% per year). Counterfactual income absent educational progress is then calculated as:

$$\tilde{y}^c = v_L \frac{y^c}{(1+r)^{\Delta S}} + v_K y^c \tag{22}$$

With y^c GDP per capita in 2019 in country c, v_L and v_K the labor and capital income shares, $\Delta S = S^{2019} - S^{1980}$ the change in average years of schooling of the working-age population, and $r = 0.1$ the return to schooling. Put simply, if average years of schooling increased by one year,

[16] The data sources and exact steps of the methodology used here differ slightly from those in Barro and Lee (2015). Barro and Lee (2015) do not report results over the 1980-2019 period, but I can compare my results to theirs for the 2001-2010 period. The two estimates are very close: education explains 15.7% of global economic growth according to their estimates, versus 19.7% according to mine. See Barro and Lee (2015), table 4.5.

then labor income would have been $\frac{1}{1.1} = 0.91$ times lower absent educational progress, while capital income would have remained unchanged.

The first line of table 6 presents the results. Mincerian returns to schooling can explain about a third of global average economic growth. The second column shows corresponding results for the global bottom 20%. Because this growth accounting decomposition relies only on cross-country data, the poorest 20% has to be defined as the poorest 20% of countries (population-weighted). Schooling gains have been relatively small for these countries. As a result, with this methodology, education can explain less than a quarter of growth for the global bottom 20%.

4.3.2. Adjusting the Labor Income Share

A first problem with this approach is that capital income shares reported in the Penn World Tables treat all mixed income as capital income. The implicit assumption is that wages are the only source of income affected by schooling; there is no return to schooling on mixed income. This does not appear to be true. OLS returns estimated in section 3.3 include mixed income and are, on average, indistinguishable from those estimated on wages only.[17] Aggregate effects of schooling estimated using the natural experiments studied in section 3.6 also include mixed income. The labor income share used for the estimation should thus include mixed income, because this is the income concept Mincerian returns to schooling apply to.

The second line of table 6 presents the results when mixed income is included in the labor income share. The Penn World Tables do not provide this decomposition, so I turn to the factor income shares estimated by Bachas et al. (2022). This adjustment alone increases the contribution of schooling to average economic growth to 43%, and its contribution to bottom 20% growth to 35%.

4.3.3. Incorporating Within-Country Inequality

In a third step, I incorporate within-country inequality in the estimation: the global poorest 20% correspond no more to the poorest 20% of countries. All other methodological ingredients stick to the standard growth accounting exercise. In particular, the average income of each income group is reduced by the same proportion within each country.

By construction, accounting for within-country inequality leaves the share of aggregate economic growth explained unchanged. However, it raises the contribution of education to bottom 20% growth from 35% to 41%, for two main reasons. First, the bottom 20% is now a composite of

[17] See appendix table G2, which compares returns to schooling by level estimated with and without including mixed income in total personal income.

individuals from low-income and middle-income countries, some of which witnessed significant schooling gains. Second, inequality has risen rapidly in many countries: growth for low-income earners has been lower than average economic growth. This second factor increases the contribution of education, simply because there is less growth among the global poor to be explained than what cross-country data suggest.

4.3.4. Incorporating Within-Country Capital Income Concentration

Fourth, I account for the fact that capital income is concentrated at the top of the distribution in each country, as discussed in section 3.5.3. For the majority of individuals belonging to the bottom 90%, almost all of income consists in wages or mixed income. The passthrough from schooling to income is thus close to 100% for low-income earners, rather than equal to the labor income share as assumed until now.

As in step 3, this does not affect the share of aggregate growth explained by education. However, it raises the contribution of education to bottom 20% growth significantly, from 41% to 51%.

4.3.5. Bringing in the Microdata

Fifth, I bring in the micro-database covering education and earnings in 150 countries collected for the purpose of this paper. I then estimate the impact of educational expansion on earnings in each country, using Mincerian returns to schooling estimated by OLS. This introduces three main differences with the previous exercise. First, returns to schooling are allowed to vary by country. Second, returns to schooling are allowed to vary by level in each country. Third, educational expansion is allowed to benefit income groups differentially in each country. For instance, expanding primary education only benefits workers who would otherwise have had no schooling, so it tends to generate more growth at the bottom than at the top of the income distribution (see section 2.4.1).

Moving from aggregate data to microdata slightly decreases the contribution of education to average growth, mainly because the average Mincerian return in my data is closer to 9% than 10%. It reduces much more strongly the contribution of education to the growth of the global bottom 20%, which falls from 51% to 39%. The main reason is that the world's poorest individuals have mostly benefited from expansions in primary and secondary education since 1980, whose average yearly returns fall well below 10%. In particular, the expansion of primary education in India has been one of the major transformations occurring during this period, with yearly returns estimated to reach less than 3% in my data (see figure 4). Heterogeneity in returns by country and level thus appears to have important effects on growth accounting estimates, implying much lower gains from schooling for the global poor than with a homogeneous 10% return.

4.3.6. Accounting for General Equilibrium Effects: Distributional Effects

Sixth, I account for supply effects redistributing income between skill groups: the expansion of education has increased the supply of skilled workers, thus depressing their wages relative to those of unskilled workers (see section 2.4.3). This step of the methodology requires data on the joint distribution of wages and education, so it can only be estimated with the microdata.

To isolate this particular distributional effect, I only adjust relative wages, leaving the average income unchanged in each country. The contribution of education to average growth is therefore the same as in the previous step. As shown in line 6 of table 6, general equilibrium effects have strongly benefited the global bottom 20%. For this group, accounting for changes in relative wages raises the contribution of education from 39% to 54%. This large impact is consistent with the empirical literature documenting important effects of changes in the supply of skilled workers on inequality and the distribution of economic growth (e.g., Goldin and Katz, 2007; Khanna, 2023; Moretti, 2004).

4.3.7. Accounting for General Equilibrium Effects: Aggregate Effects

Finally, I account for the fact that not expanding education would have been more detrimental to growth than suggested by 2019 returns. The true returns to schooling that should be used are thus higher than the returns observed in 2019 (see section 2.4.2).

This final step adjusts returns to schooling differentially by country and level, generating both aggregate and distributional effects. The contribution of education to average economic growth rises significantly, from 40% to 54%. For the bottom 20%, it increases from 54% to 69%, yielding the benchmark estimate presented at the beginning of this section.

For comparison, table 6 also reports results in which IV estimates of returns to schooling are used instead of OLS estimates. The contribution of education to average growth remains about the same, while the share of global bottom 20% growth explained rises to 76%.

4.3.8. Summary

In summary, the results presented in table 6 tell us two key facts on the role of education in reducing global poverty.

First, education explains 3 times more growth for the global poor than a standard growth accounting exercise would suggest. Changes in inequality within countries, capital income concentration, differential returns, and general equilibrium effects together imply a much more complex picture of educational expansion than that depicted by a traditional decomposition relying on aggregate data. Almost all of these additional layers of detail imply a greater contribution of education to global poverty reduction.

Second, accounting for the distributional effects of schooling within countries is of paramount importance for understanding the effect of education on global poverty reduction. Low returns to basic education imply that the contribution of schooling to global bottom 20% income growth is about 25% lower than a constant 10% return would suggest. General equilibrium effects explain over 50% of the contribution of education to global poverty reduction, by redistributing a large share of schooling gains from high-skilled to low-skilled workers.

4.4. Robustness

4.4.1. Alternative Elasticities of Substitution

My benchmark estimates assume an elasticity of substitution between skill groups of 6, at the midpoint of recent estimates focusing on long-run substitutability between skill groups. Appendix table A4 presents the share of growth explained by education by global income group under two alternative specifications. The low substitutability specification assumes elasticities of $\sigma_1 = 1.5$ (tertiary versus below tertiary), $\sigma_2 = 2.5$ (secondary versus below secondary), and $\sigma_3 = 4$ (primary versus below primary). This corresponds to the lower short-run elasticities found in the existing empirical literature (see appendix table B1). In contrast, I set $\sigma_1 = 5$, $\sigma_2 = 7$, and $\sigma_3 = 9$ in the high substitutability scenario, corresponding to elasticities at the upper bound of those found in the literature.

The low substitutability scenario implies enormous redistributive effects of educational expansion from high- to low-skill workers. As a result, the share of growth explained by education since 1980 rises from 68% to 95% for the global bottom 20%. Conversely, the high substitutability scenario implies lower supply effects, but only slightly so: the share of global bottom 20% growth explained remains as high as 60%.

4.4.2. Alternative Nesting of the CES Production Function

Another concern is that the results might be sensitive to the way the production function is specified. The specification of CES nests outlined in 3.4 is fairly standard and has been successfully used in the empirical literature (Fernández and Messina, 2018; Goldin and Katz, 2007). Yet, one may still be concerned that alternative patterns of imperfect substitutability may yield different distributional effects of educational expansion.

In appendix table A5, I consider an alternative specification in which firms first pick between workers with below and above secondary education, and then choose between subcategories of workers within these two nests.[18] With this specification, education explains 53% of global economic growth and 57% of growth for the global bottom 20%, slightly less than in the main specification but of the same order of magnitude.

4.4.3. Alternative Patterns of Schooling Expansion

Another step of the simulation has to do with identifying who benefits from private returns to schooling. In the main specification, I randomly sample individuals within age-gender cells and downgrade their education levels until matching those observed in 1980 (see section 2.4.1). While the results from the three natural experiments studied in section 3.6 suggest that this approach does a good job at capturing the distributional incidence of expansions in access to schooling, one might still be concerned about unobserved heterogeneity. For instance, if educational expansion mostly benefited children from disadvantaged socioeconomic backgrounds, one might underestimate benefits for low-income earners (who tend to come from more disadvantaged backgrounds) and overestimate aggregate effects (since individuals from more disadvantaged backgrounds might have lower expected incomes).

Extending the estimation beyond age-gender cells is not possible for the world as a whole, given the lack of data on access to schooling by socioeconomic characteristic since 1980. However, I can investigate the implications of using more refined categories for India (1983-2019), South Africa (2002-2019), and the United States (1980-2019). For India, I use historical waves of the National Sample Survey to estimate variations in educational expansion by state. For South Africa and the United States, I can similarly expand the analysis to cover differential educational progress by race and region (states in the U.S., provinces in South Africa). The results are presented in appendix figures A6, A7, and A8.[19] Using more refined categories turns out to

[18]More specifically, the production function is: $L = \left(A_L L_L^{\frac{\sigma_1-1}{\sigma_1}} + A_H L_H^{\frac{\sigma_1-1}{\sigma_1}}\right)^{\frac{\sigma_1}{\sigma_1-1}}$, with nests given by $L_L = \left(A_{non} L_{non}^{\frac{\sigma_2-1}{\sigma_2}} + A_{pri} L_{pri}^{\frac{\sigma_2-1}{\sigma_2}}\right)^{\frac{\sigma_2}{\sigma_2-1}}$ and $L_H = \left(A_{sec} L_{sec}^{\frac{\sigma_3-1}{\sigma_3}} + A_{ter} L_{ter}^{\frac{\sigma_3-1}{\sigma_3}}\right)^{\frac{\sigma_3}{\sigma_3-1}}$.

[19]These results correspond to those obtained after going through the first step of the methodology only. That is, I randomly sample individuals, downgrade educational attainment, and reduce their earnings using 2019 returns to schooling. I then plot gains from schooling as the percent difference between actual income and counterfactual income absent educational expansion.

have almost no effect on the results. It marginally reduces gains from schooling for top income earners in India and the United States, while it raises them slightly for all income groups in South Africa. I take this as reassuring evidence that the methodology adopted in this paper provides a good first-order approximation of actual distributional effects of schooling.

4.4.4. Capital Income Affected by Schooling

In my benchmark estimates, I assume that education has no effect on capital income at all, in line with standard growth accounting decompositions. This is arguably a very conservative assumption: with a constant saving rate, one should expect a fraction of schooling gains to be saved and later received in the form of capital income. There is also evidence that education can have potentially large effects on entrepreneurial income (Gennaioli et al., 2013; Queiró, 2022) or innovation (e.g., Nimier-David, 2023), both of which should translate into capital income gains. Appendix table A6 presents results in which returns to schooling are assumed to apply to both labor and capital income. This has two effects on distributional growth accounting results: it increases the contribution of education for all groups, and it increases it more for top-income groups within each country, who earn a greater fraction of income from capital. In this scenario, education explains about two-thirds of global economic growth and almost 80% of income gains for the world's poorest 20% individuals. It still explains a relatively lower fraction of growth at the top of the world distribution of income, about one-third for the top 1%, mainly because of slower educational progress in the United States and because supply effects reduce inequality within countries.

4.4.5. Education Quality

A last source of concern is that the quality of education might have changed during the period considered. This implies that changes in educational attainment since 1980 used to derive the counterfactual might not be comparable. If quality has decreased, then changes in years of schooling of the working-age population should not be valued in the same way as if it has remained the same. In the extreme case in which quality has declined sufficiently so as to fully cancel increases in quantity, one might be estimating large benefits of educational expansion even when there have been none.

I investigate trends in education quality and potential implications for the results of this paper at length in appendix D. Data on the evolution of quality are scarce, especially in the developing world. Recent test scores suggest that it has stagnated or increased in most countries (Altinok, Angrist, and Patrinos, 2018; Angrist et al., 2021), while long-run trends in literacy by cohort suggest some decline in a subset of developing economies (Le Nestour, Moscoviz, and Sandefur, 2022). All in all, there is little evidence of widespread changes in quality that would alter the main findings of this paper. Even under conservative assumptions on a potential decline in quality, I show that my main results on the world distribution of income are unaffected.

5. Education and Global Gender Inequality

This section studies the role of education in shaping the evolution of worldwide gender inequality since 1991. Sections 5.1 and 5.2 outline the conceptual framework and methodology. Section 5.3 presents the main results.

5.1. Conceptual Framework

Before moving on to the results, it is useful to conceptually distinguish the different channels through which human capital accumulation can affect gender inequality. Consider an individual i with a level of schooling s_i. They can be employed or inactive with a probability e_i, which depends on schooling: $e_i = e_i(s_i)$. When employed, their wage depends on schooling, which has a return r_i per year. Their expected income y_i is thus:

$$y_i = e_i(s_i) \cdot r_i^{s_i} \tag{23}$$

The gender gap is:

$$\begin{aligned} \Delta y &= \ln y_f - \ln y_m && (24) \\ &= \ln e_f(s_f) - \ln e_m(s_m) + (s_f \ln r_f - s_m \ln r_m) && (25) \end{aligned}$$

Where f denote women and m men. Rewriting returns to schooling for men as $r_m = \alpha r_f$:

$$\Delta y = \underbrace{\underbrace{(s_f - s_m)\ln r_f}_{\text{Differential Educational Expansion}} - \underbrace{s_m \ln \alpha}_{\text{Differential Returns to Schooling}}}_{\text{Intensive Margin}} \underbrace{+ \ln e_f(s_f) - \ln e_m(s_m)}_{\text{Extensive Margin}} \tag{26}$$

There are three main channels through which education can reduce gender inequality. First, expanding schooling differentially in the favor of women will increase their relative income. This effect will be stronger when the returns to schooling for women are high. Second, holding the distribution of educational attainment constant, the relative income of women will be greater if their returns to schooling are greater than men's. Third, schooling may have an additional impact on the gender gap by differentially affecting the labor force participation of men and women. If schooling increases the propensity of being employed, this will magnify the effect of differential educational expansion on gender inequality. I now present the data and methodology used to estimate the contribution of each of these three channels.

5.2. Methodology

To estimate the role of schooling in the reduction of global gender inequality, I apply the same methodology as the one presented in section 3. The only difference is that I allow for differential returns by gender, as well as a potential additional effect of schooling on female labor force participation. I thus construct three separate estimates of counterfactual gender inequality.

1) Differential Educational Expansion First, I consider a case in which only differential trends in educational attainment matter. As in section 3, I start by reducing education levels by age × gender cell in each country. I then reduce earnings of both men and women by the same returns to schooling, estimated over the entire working-age population.

Figure 11 shows that with the exception of Sub-Saharan Africa, there has been a significant decline in gender schooling inequalities since 1991 in all regions of the world. This reduction has been greatest in China, where the gender gap in years of schooling declined from 1.5 to 0.5. In Europe, Northern America, and Latin America, the gender education gap has reversed, with working-age women now being slightly more educated than men. Convergence in educational attainment by gender throughout the world can thus be expected to have acted as a significant driver of the reduction in gender inequality.

2) Differential Returns to Schooling Second, I incorporate differential returns to schooling conditional on being employed. To do so, I estimate Mincerian returns by gender, decomposed by education level, as in equation 16. I then reduce the earnings of men and women by gender-level-specific returns, so as to construct counterfactual earnings absent educational progress since the beginning of the period considered.

Figure 12 reproduces a well-known fact: in all regions of the world, returns to schooling are higher for women than for men (e.g., Montenegro and Patrinos, 2021).[20] This gap can be large:

[20] Appendix figure A10 extends this comparison to all countries in the dataset. Returns are higher for women in nearly all countries in the world. Of course, one may naturally be worried about selection bias. Unfortunately, the literature comparing OLS and IV estimates of differential returns to schooling by gender is scarce, although some of the studies outlined in table G4 do find comparable or higher gender gaps in returns in IV specifications.

women's returns are 2-4 percentage points higher than men's in Latin America, India, and the MENA region. Heterogeneity in returns is thus expected to amplify the inequality-reducing effects of improved access to schooling for women.

3) Extensive Margin Third, I incorporate differential effects on employment. In the benchmark specification, I run the following OLS regression in each country:

$$e_{ic} = \alpha_c + \vartheta_c \text{Educ}_{ic} \cdot \text{Gender}_{ic} + \xi_c \text{Educ}_{ic} + \digamma_c \text{Gender}_{ic} + X_{ic}\beta_c + \varepsilon_{ic} \tag{27}$$

With e_{ic} a dummy for being economically active, Educ_{ic} years of schooling, Gender_{ic} a dummy taking one for women and zero for men, and X_{ic} a vector of controls including interactions between gender and an experience quartic. The coefficient of interest is ϑ_c, capturing the differential effect of increasing education by one year on female labor force participation. This coefficient turns out to be positive in 109 countries and negative in 47 countries. In the average country, the effect of an additional year of schooling on employment is 0.7 points greater for women: increasing the average education of women by one year relative to that of men increases female labor force participation by 0.7 percentage points.

In alternative specifications, I rely on quasi-experimental evidence from a number of recent studies; their results are presented in appendix table A13. For instance, Elsayed and Shirshikova (2023) find that the staggered construction of public universities in Egypt increased female labor force participation (FLFP) for women but not for men, with an implied effect of 8 percentage points per additional year of schooling. Similarly, Keats (2018) finds that schooling increases FLPF by 7 percentage points in Uganda, using differential exposure to the 1997 elimination of primary school fees as an instrument. Two recent studies on Indonesia (Akresh, Halim, and Kleemans, 2023) and Pakistan (Khan, 2021), however, find no effect of increased access to schooling on female labor force participation. Overall, the different studies listed in table A13 find widely varying effects, ranging from 0 to 10 percentage points, with an average of about 6. Expressed as a percent increase relative to baseline, this corresponds to an average increase in FLFP of 19% per year of schooling (typically ranging from 15% to 30%).

4) Distributional Growth Accounting by Gender The final step is to estimate the contribution of education to global gender inequality reduction. This requires data on the evolution of gender inequality in each country. I rely on recent work by Neef and Robilliard (2021), who combine various sources to build the first database on the evolution of gender labor income inequality in nearly all countries in the world from 1991 to 2019. The indicator of interest is the female labor income share, defined as:

$$y_L^f(s_f, s_m) = \frac{N_f \cdot e_f(s_f) \cdot r_f^{s_f}}{\sum_{i \in \{f,m\}} N_i \cdot e_i(s_i) \cdot r_i^{s_i}} \tag{28}$$

With N_i the share of the working-age population of gender i. This indicator corresponds to the total share of labor income accruing to women, coding labor income as zero for individuals who are out of the labor force. To estimate the contribution of education to global gender inequality reduction, I thus construct counterfactual female labor income shares $\tilde{y}_L^f(\tilde{s}_f, \tilde{s}_m)$ using the methodology outlined above. I then compare the actual evolution of gender inequality to its evolution absent educational progress since 1991.

5.3. Education and Global Gender Inequality, 1991-2019

5.3.1. Education and the Global Female Labor Income Share

I start by presenting results on the contribution of schooling to gender inequality reduction from a global perspective. Table 7 compares the evolution of the global female labor income share since 1991 to its counterfactual evolution absent educational expansion. The global female labor share is calculated by taking the ratio of total labor income received by women to total labor income received by both men and women in the world as a whole.

1) Education Explains 50-80% of Global Gender Inequality Reduction There has been a decline in global gender inequality, albeit small: women received about 29% of labor income in 1991, compared to 32% in 2019.[21] The second row estimates by how much lower the female labor income share would have been if the distribution of educational attainment had remained unchanged, assuming returns to schooling are the same for men and women and no differential effect of schooling on employment. This compositional factor alone explains about half of global gender inequality reduction: the female labor income share would have increased by 1.3 points instead of the 2.8 observed. The third row incorporates heterogeneous returns to schooling by gender. Because returns to schooling are higher for women than for men in nearly all countries, this raises the contribution of education to nearly three-quarters. The last row incorporates the effect of education on female labor force participation. By this last measure, education accounts for approximately 80% of global gender inequality reduction since 1991.

[21] Appendix figure A9 compares the female labor income share in 1991 and 2019 in all countries. There has been a decline in gender inequality in 140 countries out of the 174 covered by the data. China stands out as the only large country where the female labor income share has declined.

While the magnitude of these effects might seem surprising at first, one should not forget that the dynamics of the global female labor share depend on both inequality within and between countries. Even in a world with stable gender income gaps in each country, global gender inequality can still decrease if aggregate economic growth is greater in countries with higher initial female labor income shares. The results presented in table 7 are thus capturing two separate effects of education: the effects of differential educational expansion by gender within countries, and the fact that educational attainment rose particularly rapidly in countries with lower gender inequality to begin with (such as China).

2) Education Explains 50-60% of Gender Inequality Reduction in the Average Country For the study of gender inequality, one might be more interested in understanding the particular role that education has played in reducing gender income gaps within countries. The last column of table 7 isolates this channel by presenting the average share of gender inequality reduction explained by education. To construct this indicator, I calculate the evolution of actual and counterfactual female labor income shares in each country, together with the share of gender inequality reduction explained.[22] I then take the population-weighted average of this indicator over all 150 countries covered by the data. From this angle, differential educational expansion alone accounts for 46% of gender inequality reduction in the average country. Incorporating heterogeneous returns brings this share to 57%, while accounting for employment effects raises it to almost 60%.

Robustness I investigate the sensitivity of these two main results to alternative assumptions and sample restrictions in the appendix. First, one might be worried that these two sets of results are driven by countries where gender inequality increased. In these countries, by construction, education can explain over 100% of reductions in gender inequality. China stands out as a potentially important source of concern, given its large population. Appendix table A14 reproduces table 7 after excluding China from the analysis: the results are similar. Appendix table A15 considers a more narrow restriction of the sample, excluding all 28 countries in which gender inequality declined, while still keeping countries in which education explains none of gender inequality reduction (because gender schooling gaps stagnated or increased). The contribution of education in the average country declines to about 40%.

Second, it is useful to investigate how sensitive are these results to assumptions on the effect of education on female labor force participation. Appendix table A16 compares the benchmark

[22] In countries where gender inequality increased and education mitigated this increase, I bound the share of gender inequality reduction explained by education at 100%. In countries where education increased gender inequality, I set the share of gender inequality reduction explained by education at 0%.

estimates to alternative specifications of employment effects commensurable to those found in the existing literature (see appendix table A13). I consider six specifications, in which an additional year of schooling increases employment either in absolute terms (by 4, 6, or 8 percentage points) or in relative terms (by 15%, 20%, or 25% relative to baseline employment rates). The share of global gender inequality reduction explained exceeds 100% under all these specifications, while results for the average country range from 67% to 76%.

Taken together, these findings suggest that education has been one of the most important drivers of improvements in gender labor income inequality worldwide since the 1990s. Given limited causal evidence on employment effects and differential returns to schooling by gender in each country, the results are somewhat more uncertain than those on global poverty. However, under reasonable assumptions, education accounts for at least 40-80% of the rise in the share of labor income accruing to women.

5.3.2. Education and Gender Inequality by World Region and Time Period

1) Schooling Has Reduced Gender Inequality in All Regions of the World Figures 13 and 14 extend this analysis to the evolution of gender inequality in different regions of the world. Figure 13 plots income gains for women relative to men, calculated by taking the ratio of actual to counterfactual incomes by gender in each country. Educational expansion alone has generated about 20% to 60% more growth for women than for men in all regions. Accounting for heterogeneous returns and employment effects brings this ratio to 50-150%. These last two steps have the largest effects in India and the MENA region, where returns to schooling are significantly higher for women and female labor force participation is low.[23]

Figure 14 presents the corresponding shares of gender inequality reduction explained by educational expansion since 1991. Education accounts for more than 40% of declining gender income gaps in all regions, with estimates ranging from a bit below 45% in Latin America to over 60% in the MENA region. Differential educational expansion alone can explain 20% to 50% of gender inequality reduction in all regions.

2) The Effect of Educational Expansion on Gender Inequality Has Increased Finally, I present results by world region and for the world as a whole for different time periods in appendix table A17. Two results stand out. First, the effect of education on gender inequality

[23] Appendix figure A11 plots annualized income gains from schooling for men and women in each country. In nearly all countries, schooling has generated more growth for women than for men. Accordingly, appendix figure A12 shows that the female labor income share would be significantly lower in nearly all countries in 2019 if there had been no educational expansion since 1991.

has increased over time. Income gains from schooling have been about 1.9 times greater for women since 1991 in the average country, compared to 2.4 since 2000. This rising convergence is visible in most regions, although it has been most pronounced in China and India. Second, education explains over 40% of gender inequality reduction in nearly all regions and time periods. This indicator does not display any clear time trend, given that actual gender inequality reduction has accelerated at the same time as schooling gains. Since 1991, education steadily explains about 60% of reductions in the gender income gap in the average country.

6. Discussion

This section presents a general discussion in three directions. Section 6.1 investigates what might have been the various factors explaining educational progress since 1980 and implications for the results presented in this paper. Section 6.2 discusses complementarities between education and technology and provides an empirical analysis of the role of skill-biased technical change in magnifying the growth effects of education. Section 6.3 draws on results from a companion paper to estimate the total contribution of public policies to global poverty reduction, combining direct redistribution and indirect investment benefits from education.

6.1. Where Does Schooling Come From?

In developing distributional growth *accounting*, I have done nothing else than to estimate how much lower would incomes be had education not improved, holding constant all other characteristics of the economic environment. An objection to this approach is that education might be determined by other proximate drivers of development (e.g., Hsieh and Klenow, 2010). In particular, if skill-biased technical change is a key determinant of educational expansion, separating the contribution of education loses a lot of its interest. A policymaker interested in enhancing economic growth should focus on technological progress: schooling would naturally follow. Estimating the overall contribution of technical change to human capital accumulation (and vice versa) goes far beyond this paper. In this section, I instead provide suggestive evidence that technology is unlikely to have been the dominant driver of global human capital accumulation since 1980.

1) The Global Poor Overwhelmingly Rely on Public Schools A first fact to keep in mind is that the overwhelming majority of the world's poor children are enrolled in public schools. In the 1970s and 1980s, corresponding to the period of interest to this paper, over 90% of worldwide primary school enrollment was public (World Bank, 2023). Most of the remaining 10% corresponded to children from high-income households within each country, probably putting the contribution of public schools to improved access to schooling for the global poor at over 95%.[24] In a world where governments have and continue to be the primary providers of education for poor children, the idea that market forces are the main force behind human capital accumulation since 1980 seems difficult to sustain.

2) Technological Progress Does Not Necessarily Lead to More Schooling If schooling was primarily determined by economic incentives, then one should also expect skill-biased technical change and school enrollment to closely follow each other. Yet, there are clear examples of disconnections between the two. Perhaps the most illuminating of these is the recent history of the United States, where skill-biased technical change has been exceptionally pronounced at the same time as educational progress has been among the slowest in the world. The result has been an enormous rise in wage inequality, with little indication of any endogenous adjustment in the supply of skilled workers (Autor, Goldin, and Katz, 2020; Goldin and Katz, 2008). This is not to say that market incentives do not play any role at all: there is plenty of evidence that they do.[25] The main argument is that the counterfactual world with no educational progress studied in this paper is not an impossible world. There are concrete historical examples of disconnections between education and technology that justify the use of growth accounting as a useful tool.

3) An Empirical Investigation As a last piece of complementary evidence, I investigate correlates of schooling across countries and over time. I collect data on three complementary indicators of access to schooling: expected years of schooling (corresponding to the number of years a child can hope to stay in school), net primary school enrollment, and net secondary school enrollment. I then regress these indicators on selected variables capturing different dimensions of the economic environment, from the role of government (public education spending, government effectiveness) to trade (trade-to-GDP ratio) and technology (internet usage, mobile cellular subscriptions, and the skill bias of technology estimated from the surveys compiled in this paper).

[24] India, where private schools have often been credited as the key driver of schooling expansion in recent decades, represents a particularly interesting example. Survey microdata covering the 1986-2017 period, which I have collected for a companion paper (Gethin, 2023), provide information on school attendance by type of school. There has been a large increase in the share of children enrolled in private schools over time, but this increase has been entirely driven by children from middle- and high-income households. About 85% of children coming from the poorest 20% of households were enrolled in public schools in 2017, almost the exact same share as in 1986.

[25] See for instance Atkin (2016), Blanchard and Olneyb (2017), Foster and Rosenzweig (1996), Li (2018), and Oster and Steinberg (2013).

Table A8 runs this regression across countries, for the last year available. Table A9 extends this analysis to the 1980-2019 period with country and year fixed effects. The takeaway is that public education expenditure stands out as the only variable robustly correlated with schooling in both cross-section and panel data. The skill bias of technology is not significantly related to schooling across countries, and if anything enters the regression with the wrong sign. There is also no clear evidence that countries with greater economic growth or faster adoptions of new technologies have had larger increases in schooling since 1980 (if anything, the opposite seems to be true). In contrast, real public education expenditure per child is strongly associated with improvements in access to schooling across all three indicators. These results should of course not be interpreted causally, but they suggest that technological progress is unlikely to have been behind the major schooling expansions of the past decades studied in this paper.

6.2. Education and Skill-Biased Technical Change

The main results of this paper are based on building counterfactual income distributions in 2019, bringing education levels back to their 1980 value. This amounts to answering the following question: how much lower would incomes be if education had not improved, holding all other factors to their 2019 values? An alternative way of estimating gains from schooling would instead be to use surveys from the 1980s to estimate the effect of increasing education levels to their 2019 value. This amounts to answering a slightly different question: how much higher would incomes be if education had improved, holding all other factors to their 1980 values? If we stick to the CES production function with labor-augmenting technology terms, the difference between these two estimates turns out to provide useful insights into the role of skill-biased technical change in amplifying the growth effects of schooling.

6.2.1. Backward Accounting: Schooling With Relative Efficiency Gains

The main results of this paper correspond to what one might call "backward accounting," bringing education levels back to their 1980 value. Formally, let $A^t = \frac{A_H^t}{A_L^t}$ be the skill bias of technology (the relative efficiency of skilled workers) and L^t the vector of skill supplies at time t. Output is a function of both: $Y^t = Y(A^t, L^t)$. The main exercise of this paper amounts to constructing:

$$\tilde{Y}^{2019}_{\text{backward}} = Y(A^{2019}, L^{1980}) \tag{29}$$

The share of growth explained by education is:

$$\text{Success}_{\text{backward}} = 1 - \frac{\frac{Y(A^{2019}, L^{1980})}{Y(A^{1980}, L^{1980})} - 1}{\frac{Y(A^{2019}, L^{2019})}{Y(A^{1980}, L^{1980})} - 1} \tag{30}$$

The estimation thus amounts to comparing growth rates with and without educational expansion, *assuming that the relative efficiency of skilled workers evolved as it did*, from A^{1980} to A^{2019}. In this model, not expanding education has large negative effects on growth and inequality. In particular, supply effects magnify the growth effects of education, because the loss in income from not expanding education would have been greater than what 2019 returns suggest.

6.2.2. Forward Accounting: Schooling Without Relative Efficiency Gains

An alternative would be to work with "forward accounting," taking incomes in 1980 and estimating the growth effects of moving to 2019 education levels. This amounts to constructing:

$$\tilde{Y}^{2019}_{\text{backward}} = Y(A^{1980}, L^{2019}) \tag{31}$$

The share of growth explained by education is:

$$\text{Success}_{\text{forward}} = \frac{\frac{Y(A^{1980}, L^{2019})}{Y(A^{1980}, L^{1980})} - 1}{\frac{Y(A^{2019}, L^{2019})}{Y(A^{1980}, L^{1980})} - 1} \tag{32}$$

The interpretation is now different: it amounts to comparing actual growth to a world in which only education would have increased, while technology would have remained to its 1980 value. Supply effects now decrease the contribution of education to aggregate growth: the return to schooling declines as education expands when the skill bias of technology does not.

If relative efficiency A has not changed, then the two estimates should be identical. Indeed, because the supply of skilled workers was lower in 1980, the return to schooling should be much higher in 1980 than in 2019. Aggregate gains from schooling estimated by increasing incomes in 1980 (using 1980 returns diminished by supply effects) should then be equal to those estimated by reducing incomes in 2019 (using 2019 returns augmented by supply effects).

If relative efficiency has changed, however, the two estimates will differ. In the presence of skill-biased technical change ($A^{2019} > A^{1980}$), in particular, we should expect backward accounting to deliver greater growth effects of education: $\text{Success}_{\text{backward}} > \text{Success}_{\text{forward}}$. The reason is intuitive: in a world with rising relative efficiency of skilled workers, expanding schooling turns out to be a much more profitable investment than in a world where it remains unchanged. Comparing estimates of backward and forward growth accounting across countries can then tell us something about the role of skill-biased technical change in enhancing the growth effects of educational expansion.

6.2.3. An Empirical Investigation

Unfortunately, 1980s survey data are available for almost none of the countries covered by this paper. However, I was able to find and harmonize comparable surveys covering personal income and education in the early 2000s for 33 countries: 14 European countries, 14 Latin American countries, the United States, Indonesia, Thailand, Ghana, and South Africa. I can thus investigate how estimates from backward and forward accounting differ across these countries for the 2000-2019 period. Surveys fielded at the beginning and end of the period are not always perfectly comparable, so the results should be interpreted with some care, but they can still provide suggestive evidence.

I estimate backward accounting results using the same methodology as in the rest of the paper. For forward accounting, I take 2000 surveys and increase education levels to their 2019 values in each country. I then estimate returns to schooling in 2000 and reduce them to reach true returns to schooling, using the exact opposite step as the one outlined in section 2.2. Finally, I estimate distributional effects by adjusting relative wages as in the rest of the paper.

6.2.4. Results

Appendix table A10 compares estimated income gains from schooling under backward (with efficiency gains) and forward (without efficiency gains) accounting, separately for the population as a whole and the bottom 50%.

Gains from schooling are almost always lower under forward than backward growth accounting. This is consistent with significant increases in the relative efficiency of skilled workers, which have made the benefits of expanding access to schooling much larger than an analysis holding technology fixed would suggest (which is what forward accounting does). The gap between the two estimates can be particularly large for the bottom 50% of earners. Indeed, skill-biased technical change has both aggregate and distributional effects. When the skill bias of technology is rising rapidly, expanding access to schooling becomes a particularly powerful tool for ensuring that the benefits of technological progress are broadly shared. This is just an example of the "race between education and technology" (Goldin and Katz, 2008).

There are significant variations across countries. Europe stands out as the region where the gap between the two estimates is the largest: for the bottom 50%, they differ by a factor of 4. This is consistent with the fact that skill-biased technical change has been particularly pronounced in high-income countries in recent decades (this does not stand out in the case of the United States mainly because educational progress was particularly small during this period). In contrast, forward and backward accounting yield almost identical results in Brazil and Mexico, in line with recent evidence pointing to stagnating or even declining demand for skilled labor in these countries (Fernández and Messina, 2018).

Appendix table A11 extends this analysis to the share of growth explained by education. Forward accounting typically explains 25-30% less growth than backward accounting, with significant variations across countries. For the bottom 50%, forward accounting still explains an important fraction of growth in most countries covered in this analysis, including about 60% in Europe, 50% in Indonesia, and over 100% in Brazil and Mexico. Even absent skill-biased technical change, expanding education would still have had significant effects on inequality and growth for the poorest individuals in these regions of the world.

In summary, education and labor-augmenting technology act as strong complements. Skill-biased technology without education will generate large increases in inequality with little growth for low-income earners. A perfect example is the trajectory of the United States since 1980, where pretax incomes have literally stagnated for the bottom 50% of earners despite strong macroeconomic growth (Piketty, Saez, and Zucman, 2018). Schooling expansions without technology will reduce inequality but will display lower and decreasing returns (as was perhaps visible in Latin America in the 2000s). With labor-augmenting technology and imperfect substitution, the classic separation between education and total factor productivity becomes more difficult to conceptualize.[26] Arguably, the historical contribution of education to economic growth should be evaluated in light of how skill-biased technical change has actually evolved. This is what I have attempted to do in this paper.

6.3. An Estimate of the Total Contribution of Public Policies to Global Poverty Reduction

I conclude this paper with a combined analysis of direct effects of government transfers on poverty and indirect effects of education on pretax incomes. Public policies contribute to reducing poverty through two channels at a given point in time. First, individuals benefit from cash transfers and in-kind transfers that increase their posttax incomes. In a companion

[26] Caselli and Ciccone (2019) and Jones (2019) provide an interesting discussion along these lines.

paper, I show that such direct government redistribution can account for 30% of global poverty reduction since 1980 (Gethin, 2023). Second, public policies can also contribute to increasing future pretax earnings; studying this indirect contribution was the objective of this paper, with a focus on education. Putting these two estimates together can then give us an approximate estimate of the total contribution of public policies to global poverty reduction. This estimate is arguably partial, given that public education is not the only type of policy contributing to pretax earnings growth, but it can at least be viewed as a lower bound.

Table 8 compares the evolution of global poverty, the average income of the world's poorest 20%, and the average income of the world's poorest 50% before and after accounting for direct redistribution and indirect benefits from education. Absent educational expansion since 1980, poverty would have declined by about 32% in terms of pretax income, compared to an actual reduction of 55%. Removing taxes from individual incomes and adding government transfers yields an estimate of global poverty in terms of posttax income, which declined by 70%. By this measure, direct redistribution and indirect benefits from education together account for about 55% (1 - $\frac{32}{70}$) of global poverty reduction since 1980. Using a similar reasoning, public policies account for about 80% of real income gains for the world's poorest 20% individuals, and about two-thirds of gains for the poorest half of the world's population.

As a robustness check, I reproduce this analysis using World Bank data instead of data from the World Inequality Database. The results are presented in appendix table A12. With this data source, public policies account for about 46% of global poverty reduction since 1980, slightly less than with the WID data. This is not surprising, given that poverty at this threshold already declined by almost 80% in terms of pretax income, which mechanically puts an upper bound on the role played by direct redistribution. Public policies explain close to 70% of income gains for the global bottom 20% and bottom 50%, corresponding to the same orders of magnitude as with the WID data. The takeaway is that public policies can account for 50-80% of the reduction of global poverty since 1980, and potentially much more if we were to account for the indirect growth effects of public healthcare, transport infrastructure, housing policies, public order and safety, and increasing investments in other public goods observed in the past decades.

7. Conclusion

This article represented a first attempt at estimating the role played by education in the historical reduction of global poverty. Combining a stylized model of education and the wage structure with tools borrowed from growth accounting, I proposed a "distributional growth accounting" framework identifying the contribution of education to real income growth within and across countries. Under conservative assumptions, private returns to schooling can explain a large fraction of real income gains among the world's poorest individuals, in the order of 60-70% and potentially more. It can also account for over half of the rise in the share of income accruing to women. This puts public education policies at the center of the remarkable reduction of poverty and gender inequality observed in the past decades.

The focus of this article was on private returns to schooling, yet much remains to be done on other dimensions of human capital such as work experience, indirect effects of education on technology, and human capital externalities. How has educational expansion contributed to innovation and its diffusion worldwide? To what extent has population ageing affected economic growth through returns to experience, and how do these effects vary across skill groups in developed and developing economies? Understanding these important economic questions would allow for a fuller understanding of the role played by human capital in shaping global inequalities.

More generally, the results presented in this article call for further research on the structural drivers of changes in the world distribution of income since 1980. The approach adopted in this paper could be extended to many other key transformations of the past decades, including globalization, structural change, financialization, capital accumulation, population ageing, democratization, and changing gender norms. The microeconomics literature provides ample empirical evidence on the economic effects of these factors in specific contexts. Combined with new data collection efforts and adequate theoretical frameworks, this evidence could be aggregated to shed light on the role played by these long-run processes in the reduction of global poverty.

References

Acemoglu, Daron (1998). "Why do new technologies complement skills? Directed technical change and wage inequality". In: *The Quarterly Journal of Economics* 113.1, pp. 1055–1089.

Acemoglu, Daron and Joshua Angrist (2000). "How Large Are Human-Capital Externalities? Evidence from Compulsory Schooling Laws". In: *NBER Macroeconomics Annual* 15.1, pp. 9–59.

Acemoglu, Daron and David Autor (2011). "Skills, Tasks and Technologies: Implications for Employment and Earnings". In: *Handbook of Labor Economics, Volume 4B*. Ed. by Orley Ashenfelter and David Card. North Holland, pp. 1043–1171.

Akresh, Richard, Daniel Halim, and Marieke Kleemans (2023). "Long-Term and Intergenerational Effects of Education: Evidence from School Construction in Indonesia". In: *The Economic Journal* 133.650, pp. 582–612.

Altinok, Nadir, Noam Angrist, and Harry A. Patrinos (2018). “Global Data Set on Education Quality (1965-2015)”. World Bank Policy Research Working Paper No. 8314.

Angrist, Joshua D. (1995). “The Economic Returns to Schooling in the West Bank and Gaza Strip”. In: *The American Economic Review* 85.5, pp. 1065–1087.

Angrist, Noam, Simeon Djankov, Pinelopi K. Goldberg, and Harry A. Patrinos (2021). “Measuring human capital using global learning data”. In: *Nature* 592.1, pp. 403–408.

Assaad, Ragui, Abdurrahman B. Aydemir, Meltem Dayıoglu, and Murat Güray Kırdar (2023). “Wage Returns to Human Capital Resulting from an Extra Year of Primary School: Evidence from Egypt”. IZA Discussion Paper.

Atkin, David (2016). “Endogenous Skill Acquisition and Export Manufacturing in Mexico”. In: *American Economic Review* 106.8, 2046–2085.

Autor, David, Claudia Goldin, and Lawrence F. Katz (2020). “Extending the Race between Education and Technology”. In: *AEA Papers and Proceedings* 110.1, pp. 347–351.

Bachas, Pierre, Matthew H. Fisher-Post, Anders Jensen, and Gabriel Zucman (2022). “Globalization and Factor Income Taxation”. NBER Working Paper 29819.

Barro, Robert J. and Jong-Wha Lee (2013). “A new data set of educational attainment in the world, 1950–2010”. In: *Journal of Development Economics* 104.1, pp. 184–198.

– (2015). *Education Matters: Global Schoolig Gains from the 19th Century to the 21st Century*. Oxford University Press.

Bick, Alexander, Nicola Fuchs-Schündeln, and David Lagakos (2018). “How Do Hours Worked Vary With Income? Cross-Country Evidence and Implications”. In: *American Economic Review* 108.1, pp. 170–199.

Bils, Mark, Baris Kaymak, and Kai-Jie Wu (2022). “Labor Substitutability Among Schooling Groups”. NBER Working Paper.

Blanchard, Emily J. and William W. Olneyb (2017). “Globalization and human capital investment: Export composition drives educational attainment”. In: *Journal of International Economics* 106.1, pp. 165–183.

Blanchet, Thomas, Lucas Chancel, and Amory Gethin (2022). “Why Is Europe More Equal Than the United States?” In: *American Economic Journal: Applied Economics* 14.4, pp. 1–40.

Blanchet, Thomas, Ignacio Flores, and Marc Morgan (2022). “The weight of the rich: improving surveys using tax data”. In: *The Journal of Economic Inequality* 20.1, pp. 119–150.

Bowlus, Audra, Lance Lochner, Chris Robinson, and Eda Suleymanoglu (2021). “Wages, Skills, and Skill-Biased Technical Change: The Canonical Model Revisited”. Working Paper.

Brunello, Giorgio, Guglielmo Weber, and Christoph T. Weiss (2015). “Books are Forever: Early Life Conditions, Education and Lifetime Earnings in Europe”. In: *The Economic Journal* 127.600, pp. 271–296.

Calonico, Sebastian, Matias D. Cattaneo, and Rocío Titiunik (2014). "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs". In: *Econometrica* 82.6, 2295–2326.

Card, David (1999). "The Causal Effect of Education on Earnings". In: *Handbook of Labor Economics*. Ed. by Orley C. Ashenfelter and David Card. Elsevier, pp. 1801–1863.

– (2001). "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems". In: *Econometrica* 69.5, pp. 1127–1160.

Carneiro, Pedro, Michael Lokshin, and Nithin Umapathi (2017). "Average and Marginal Returns to Upper Secondary Schooling in Indonesia". In: *Journal of Applied Econometrics* 32.1, pp. 16–36.

Caselli, Francesco and Antonio Ciccone (2013). "The contribution of schooling in development accounting: Results from a nonparametric upper bound". In: *Journal of Development Economics* 104.1, pp. 199–211.

– (2019). "The Human Capital Stock: A Generalized Approach: Comment". In: *American Economic Review* 109.3, pp. 1155–1174.

Caselli, Francesco and Wilbur John II Coleman (2006). "The World Technology Frontier". In: *American Economic Review* 96.3, pp. 499–522.

Chancel, Lucas and Thomas Piketty (2021). "Global Income Inequality, 1820–2020: the Persistence and Mutation of Extreme Inequality". In: *Journal of the European Economic Association* 19.6, 3025–3062.

Chancel, Lucas, Thomas Piketty, Emmanuel Saez, and Gabriel Zucman (2022). *World Inequality Report 2022*. Harvard University Press.

Chatterjee, Aroop, Léo Czajka, and Amory Gethin (2021). "Can Redistribution Keep Up with Inequality? Evidence from South Africa, 1993-2019". Working Paper.

– (2022). "Wealth Inequality in South Africa, 1993-2017". In: *World Bank Economic Review* 36.1, pp. 19–36.

Chauvin, Juan Pablo, Edward Glaeser, Yueran Ma, and Kristina Tobio (2017). "What is different about urbanization in rich and poor countries? Cities in Brazil, China, India and the United States". In: *Journal of Urban Economics* 98.1, pp. 17–49.

Che, Yi and Lei Zhang (2018). "Human Capital, Technology Adoption and Firm Performance: Impacts of China's Higher Education Expansion in the Late 1990s". In: *The Economic Journal* 128.614, pp. 2282–2320.

Chen, Shaohua and Martin Ravallion (2010). "The Developing World Is Poorer Than We Thought, But No Less Successful In the Fight Against Poverty". In: *Quarterly Journal of Economics* 125.4, pp. 1577–1625.

Chicoine, Luke (2021). "Free Primary Education, Fertility, and Women's Access to the Labor Market: Evidence from Ethiopia". In: *The World Bank Economic Review* 35.2, 480–498.

Ciccone, Antonio and Giovanni Peri (2005). “Long-Run Substitutability Between More and Less Educated Workers: Evidence from U.S. States, 1950-1990”. In: *The Review of Economics and Statistics* 87.4, pp. 652–663.

– (2006). “Identifying Human-Capital Externalities: Theory with Applications”. In: *Review of Economic Studies* 73.1, pp. 381–412.

Clay, Karen, Jeff Lingwall, and Melvin Jr. Stephens (2021). “Laws, educational outcomes, and returns to schooling evidence from the first wave of U.S. state compulsory attendance laws”. In: *Labour Economics* 68.1, pp. 1019–1035.

Collin, Matthew and David Weil (2020). “The Effect of Increasing Human Capital Investment on Economic Growth and Poverty: A Simulation Exercise”. In: *Journal of Human Capital* 14.1, pp. 43–83.

Cui, Ying, Hong Liu, and Liqiu Zhao (2019). “Mother’s education and child development: Evidence from the compulsory school reform in China”. In: *Journal of Comparative Economics* 47.1, pp. 669–692.

De Rosa, Mauricio, Ignacio Flores, and Marc Morgan (2022). “More Unequal or Not as Rich? Revisiting the Latin American exception”. Stone Center Working Paper.

Delesalle, Esther (2021). “The effect of the Universal Primary Education program on consumption and on the employment sector: Evidence from Tanzania”. In: *World Development* 142.1, pp. 1053–1070.

Deming, David J. (2022). “Four Facts about Human Capital”. In: *Journal of Economic Perspectives* 36.3, pp. 75–102.

– (2023). “Multidimensional Human Capital and the Wage Structure”. NBER Working Paper.

Duflo, Esther (2001). “Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment”. In: *American Economic Review* 91.4, pp. 795–813.

– (2004). “The medium run effects of educational expansion: evidence from a large school construction program in Indonesia”. In: *Journal of Development Economics* 74.1, pp. 163–197.

Ekhator-Mobayode, Uche E. and Johannes Hoogeveen (2022). “Microdata collection and openness in the Middle East and North Africa”. In: *Data and Policy* 4.31, pp. 1–17.

Elsayed, Ahmed and Alina Shirshikova (2023). “The women-empowering effect of higher education”. In: *Journal of Development Economics* 163.1, pp. 1031–1041.

Erten, Bilge and Pinar Keskin (2018). “For Better or for Worse?: Education and the Prevalence of Domestic Violence in Turkey”. In: *American Economic Journal: Applied Economics* 10.1, pp. 64–105.

Fang, Hai, Karen N. Eggleston, John A. Rizzo, Scott Rozelle, and Richard J. Zeckhauser (2012). "The Returns to Education in China: Evidence from the 1986 Compulsory Education Law". NBER Working Paper.

Fernández, Manuel and Julián Messina (2018). "Skill premium, labor supply, and changes in the structure of wages in Latin America". In: *Journal of Development Economics* 135.1, pp. 555–573.

Foster, Andrew D. and Mark R. Rosenzweig (1996). "Technical Change and Human-Capital Returns and Investments: Evidence from the Green Revolution". In: *American Economic Review* 86.4, pp. 931–953.

Gennaioli, Nicola, Rafael La Porta, Florencio Lopez-de Silanes, and Andrei Shleifer (2013). "Human Capital and Regional Development". In: *The Quarterly Journal of Economics* 128.1, 105–164.

Gethin, Amory (2022). "Who Benefits from Public Goods? Evidence from South Africa". Working Paper.

– (2023). "Revisiting Global Poverty Reduction: Public Goods and the World Distribution of Income, 1980-2019". Working Paper: https://amory-gethin.fr/files/pdf/Gethin2023PublicGoods.pdf.

Glaeser, Edward and Ming Lu (2018). "Human-Capital Externalities in China". NBER Working Paper.

Goldin, Claudia and Lawrence F. Katz (2007). "The Race between Education and Technology: The Evolution of U.S. Educational Wage Differentials, 1890 to 2005". NBER Working Paper.

– (2008). *The Race Between Education and Technology*. Harvard University Press.

Grépin, Karen A. and Bharadwaj Prashant (2015). "Maternal education and child mortality in Zimbabwe". In: *Journal of Health Economics* 44.1, 97–117.

Gunderson, Morley and Philip Oreopolous (2020). "Returns to education in developed countries". In: *The Economics of Education*. Ed. by Steve Bradley and Colin Green. Academic Press, pp. 39–51.

Guo, Junjie, Nicolas Roys, and Ananth Seshadri (2018). "Estimating Aggregate Human Capital Externalities". Working Paper.

Hall, Robert E. and Charles I. Jones (1999). "Why Do Some Countries Produce So Much More Output Per Worker Than Others?" In: *The Quarterly Journal of Economics* 114.1, pp. 83–116.

Hanushek, Eric A. and Dennis D. Kimko (2000). "Schooling, Labor-Force Quality, and the Growth of Nations". In: *American Economic Review* 90.5, pp. 1184–1208.

Hanushek, Eric A., Jens Ruhose, and Ludger Woessman (2017). "Knowledge Capital and Aggregate Income Differences: Development Accounting for US States". In: *American Economic Journal: Macroeconomics* 9.4, pp. 184–224.

Heckman, James J., John Eric Humphries, and Gregory Veramendi (2018). “Returns to Education: The Causal Effects of Education on Earnings, Health, and Smoking”. In: *Journal of Political Economy* 126.1, pp. 197–246.

Heckman, James J., Lance Lochner, and Christopher Taber (1998). “Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents”. In: *Review of Economic Dynamics* 1.1, pp. 1–58.

Hendricks, Lutz and Todd Schoellman (2018). “Human Capital and Development Accounting: New Evidence from Wage Gains at Migration”. In: *The Quarterly Journal of Economics* 133.2, 665–700.

– (2023). “Skilled Labor Productivity and Cross-Country Income Differences”. In: *American Economic Journal: Macroeconomics* 15.1, pp. 240–268.

Hermo, Santiago, Miika Päällysaho, David Seim, and Jesse Shapiro (2022). “Labor Market Returns and the Evolution of Cognitive Skills: Theory and Evidence”. In: *Quarterly Journal of Economics* 137.4, pp. 2309–2361.

Hershbein, Brad, Melissa S. Kearney, and Luke W. Pardue (2020). “College Attainment, Income Inequality, and Economic Security: A Simulation Exercise”. In: *AEA Papers and Proceedings* 110.1, pp. 352–355.

Hicks, Daniel and Huiqiong Duanc (2023). “Education as opportunity? The causal effect of education on labor market outcomes in Jordan”. In: *Oxford Development Studies* 51.2, 179–197.

Hsiao, Allan (2023). “Educational Investment in Spatial Equilibrium: Evidence from Indonesia”. Working Paper.

Hsieh, Chang-Tai and Peter J. Klenow (2010). “Development Accounting”. In: *American Economic Journal: Macroeconomics* 2.1, pp. 207–223.

Huang, Bin and Yu Zhu (2022). “Returns to education in China: Evidence from the great higher education expansion”. In: *China Economic Review* 74, pp. 1018–1104.

Imbens, Guido W. and Karthik Kalyanaraman (2012). “Optimal Bandwidth Choice for the Regression Discontinuity Estimator”. In: *Review of Economic Studies* 79.3, pp. 933–959.

Iranzo, Susana and Giovanni Peri (2009). “Schooling Externalities, Technology, and Productivity: Theory and Evidence from U.S. States”. In: *The Review of Economics and Statistics* 91.2, pp. 420–431.

Jones, Benjamin F. (2014). “The Human Capital Stock: A Generalized Approach”. In: *American Economic Review* 104.11, 3752–3777.

– (2019). “The Human Capital Stock: A Generalized Approach: Reply”. In: *American Economic Review* 109.3, pp. 1175–1195.

Karachiwalla, Naureen and Giordano Palloni (2019). “Human Capital and Structural Transformation: Quasi-Experimental Evidence from Indonesia”. Working Paper.

Katz, Lawrence F. and Kevin M. Murphy (1992). “Changes in Relative Wages, 1963-1987: Supply and Demand Factors”. In: *The Quarterly Journal of Economics* 107.1, pp. 35–78.

Keats, Anthony (2018). “Women’s schooling, fertility, and child health outcomes: Evidence from Uganda’s free primary education program”. In: *Journal of Development Economics* 135.1, pp. 142–159.

Khan, Salman (2021). “Unmet Expectations: The Impacts of School Construction on Female Outcomes in Rural Punjab, Pakistan”. Working Paper.

Khanna, Gaurav (2023). “Large-Scale Education Reform in General Equilibrium: Regression Discontinuity Evidence from India”. In: *Journal of Political Economy* 131.2, pp. 549–591.

Kim, Dahye (2023). “Women’s Higher Education, Assortative Mating, and Empowerment: Long-Term Evidence from College Enrollment Expansion in South Korea”. In: *The Journal of Development Studies* 59.3, pp. 430–447.

Kleven, Henrik and Camille Landais (2017). “Gender Inequality and Economic Development: Fertility, Education, and Norms”. In: *Economica* 84.334, pp. 180–209.

Kyui, Natalia (2016). “Expansion of higher education, employment and wages: Evidence from the Russian Transition”. In: *Labour Economics* 39.1, pp. 68–87.

Lakner, Christoph and Branko Milanovic (2016). “Global Income Distribution: From the Fall of the Berlin Wall to the Great Recession”. In: *The World Bank Economic Review* 30.2, pp. 203–232.

Le Nestour, Alexis, Laura Moscoviz, and Justin Sandefur (2022). “The Long-Run Decline of Education Quality in the Developing World”. Working Paper.

Lemieux, Thomas (2006). “The "Mincer Equation" Thirty Years After Schooling, Experience, and Earnings”. In: *Jacob Mincer. A Pioneer of Modern Labor Economics*. Ed. by Shoshana Grossbard. Springer. Chap. 11, 127–145.

Lemieux, Thomas and David Card (2001). “Education, Earnings, and the ’Canadian G.I. Bill’”. In: *The Canadian Journal of Economics / Revue canadienne d’Economique* 34.2, pp. 313–344.

Li, Bingjing (2018). “Export expansion, skill acquisition and industry specialization: evidence from China”. In: *Journal of International Economics* 114.1, pp. 346–361.

Mankiw, Gregory N., David Romer, and David N. Weil (1992). “A Contribution to the Empirics of Economic Growth”. In: *Quarterly Journal of Economics* 107.2, pp. 407–437.

Montenegro, Claudio E. and Harry Anthony Patrinos (2021). “A data set of comparable estimates of the private rate of return to schooling in the world, 1970–2014”. In: *International Journal of Manpower* 1.1, pp. 1–20.

Moretti, Enrico (2004). “Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data”. In: *Journal of Econometrics* 121, pp. 175–212.

Morrisson, Christian and Fabrice Murtin (2013). “The Kuznets curve of human capital inequality: 1870–2010”. In: *The Journal of Economic Inequality* 11.1, 283–301.
Murphy, Kevin M., W. Craig Riddell, and Paul M. Romer (1998). “Wages, Skills, and Technology in the United States and Canada”. NBER Working Paper.
Navarro-Sola, Laia (2021). “Secondary Schools with Televised Lessons: The Labor Market Returns of the Mexican Telesecundaria”. HCEO Working paper.
Neef, Theresa and Anne-Sophie Robilliard (2021). “Half the Sky? The Female Labor Income Share in a Global Perspective”. In: *Journal of Comparative Economics* 47.1, pp. 669–692.
Nimier-David, Elio (2023). “Local Human Capital and Firm Creation: Evidence from the Massification of Higher Education in France”. Working Paper.
Oliobi, Ifeatu (2022). “Access to Education and Family Formation: Evidence from University Expansion in Nigeria”. Working Paper.
Oster, Emily and Bryce Millett Steinberg (2013). “Do IT service centers promote school enrollment? Evidence from India”. In: *Journal of Development Economics* 104.1, 123–135.
Oyelere, Ruth Uwaifo (2010). “Africa’s education enigma? The Nigerian story”. In: *Journal of Development Economics* 91.1, pp. 128–139.
Piketty, Thomas, Emmanuel Saez, and Gabriel Zucman (2018). “Distributional National Accounts: Methods and Estimates for the United States”. In: *The Quarterly Journal of Economics* 133.2, pp. 553–609.
Pinkovskiy, Maxim and Xavier Sala-i-Martin (2016). “Lights, Camera … Income! Illuminating the National Accounts-Household Surveys Debate”. In: *The Quarterly Journal of Economics* 131.2, pp. 579–631.
Queiró, Francisco (2022). “Entrepreneurial Human Capital and Firm Dynamics”. In: *The Review of Economic Studies* 89.4, 2061–2100.
Roodman, David (2022). “Schooling and Labor Market Consequences of School Construction in Indonesia: Comment”. Working Paper.
Rossi, Federico (2020). “Human Capital and Macroeconomic Development: A Review of the Evidence”. In: *The World Bank Research Observer* 35.2, pp. 227–262.
– (2022). “The Relative Efficiency of Skilled Labor across Countries: Measurement and Interpretation”. In: *American Economic Review* 112.1, pp. 235–266.
Sakellariou, Chris (2006). “Education policy reform, local average treatment effect and returns to schooling from instrumental variables in the Philippines”. In: *Applied Economics* 38.4, pp. 473–481.
Sala-i-Martin, Xavier (2006). “The World Distribution of Income: Falling Poverty and... Convergence, Period”. In: *The Quarterly Journal of Economics* 121.2, pp. 351–397.

Schoellman, Todd (2012). “Education Quality and Development Accounting”. In: *The Review of Economic Studies* 79.1, 388–417.

Solow, Robert M. (1957). “Technical Change and the Aggregate Production Function”. In: *Review of Economics and Statistics* 39.1, pp. 312–320.

Spohr, Chris A. (2003). “Formal schooling and workforce participation in a rapidly developing economy: evidence from compulsory junior high school in Taiwan”. In: *Journal of Development Economics* 70.1, pp. 291–327.

Stephens, Melvin Jr. and Dou-Yan Yang (2014). “Compulsory Education and the Benefits of Schooling”. In: *American Economic Review* 104.6, pp. 1777–1792.

United Nations (2023). *The Sustainable Development Goals Report*. United Nations.

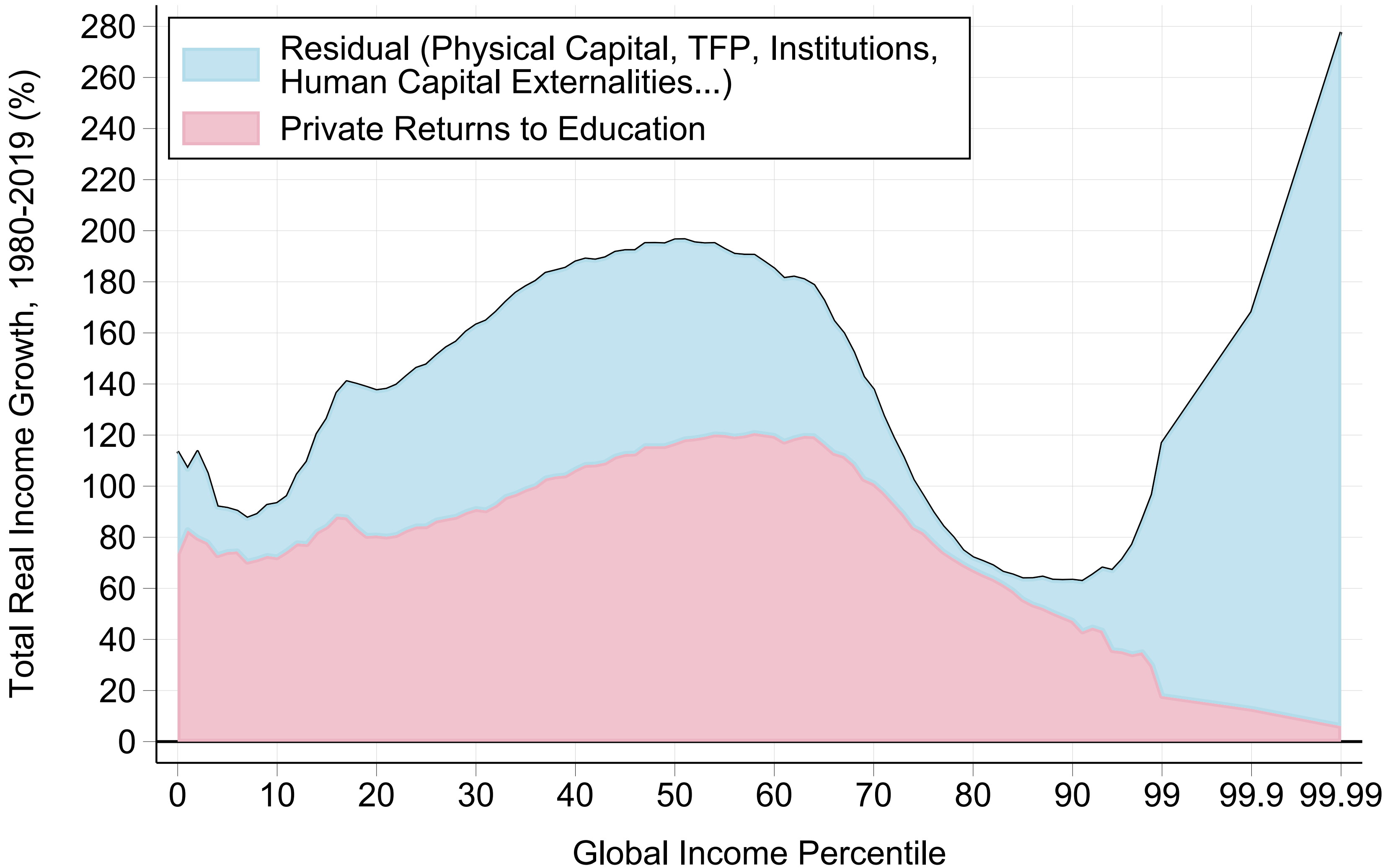
Vu, Khoa and Tu-Anh Vu-Thanh (2022). “Higher Education Expansion, Labor Market, and Firm Productivity in Vietnam”. Working Paper.

Wantchekon, Leonard, Marko Klašnja, and Natalija Novta (2015). “Education and Human Capital Externalities: Evidence from Colonial Benin”. In: *The Quarterly Journal of Economics* 130.2, 703–757.

World Bank (2023). *World Development Indicators*. World Bank.

Zimmerman, Seth D. (2014). “The Returns to College Admission for Academically Marginal Students”. In: *Journal of Labor Economics* 32.4, pp. 711–754.

Figure 1 – Education and the Distribution of Global Economic Growth, 1980-2019



Notes. The figure plots total real income growth by global income percentile from 1980 to 2019, decomposing it into a part that can be explained by private returns to schooling and an unexplained component. The upper shaded area represents the growth rates that would have prevailed absent any improvement in the education of the world's working-age population since 1980. The lower shaded area represents the corresponding contribution of education to economic growth. Taking the ratio between this contribution and actual growth rates, education explains about 60-80% of growth for the world's 20% poorest individuals. The income concept is pretax income per capita.

Figure 2 – Survey Microdata Coverage

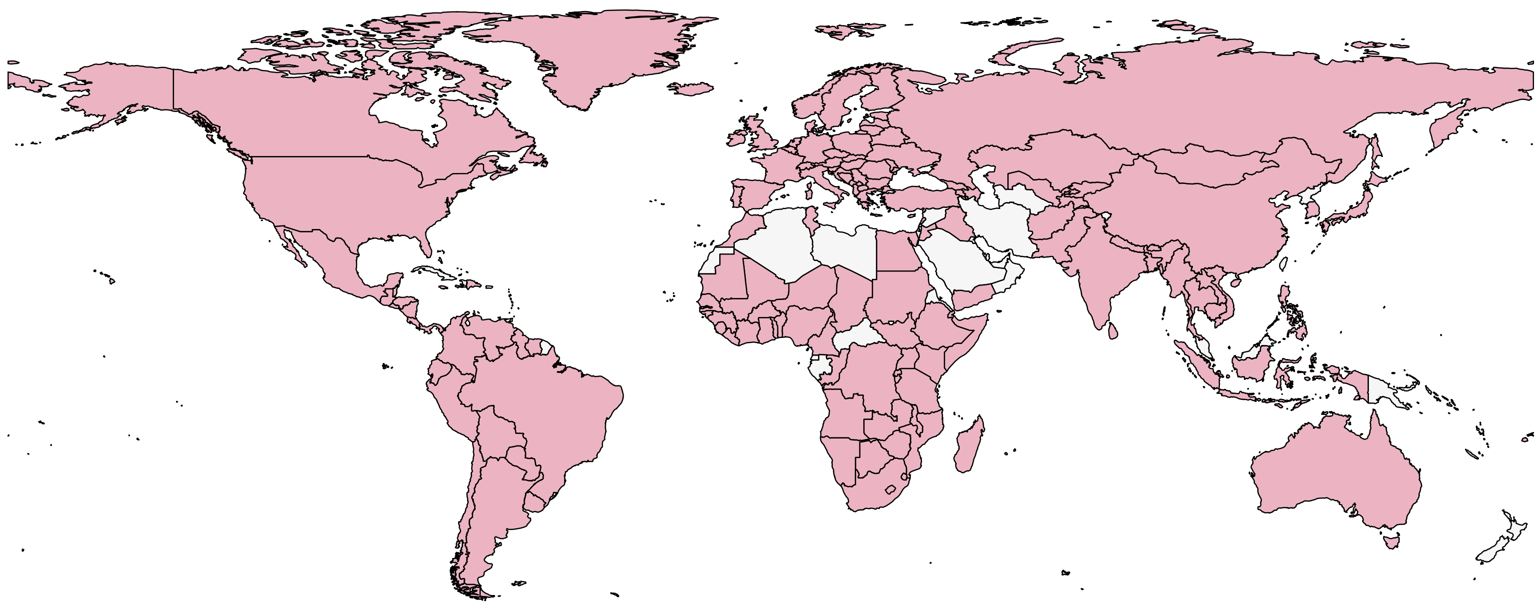
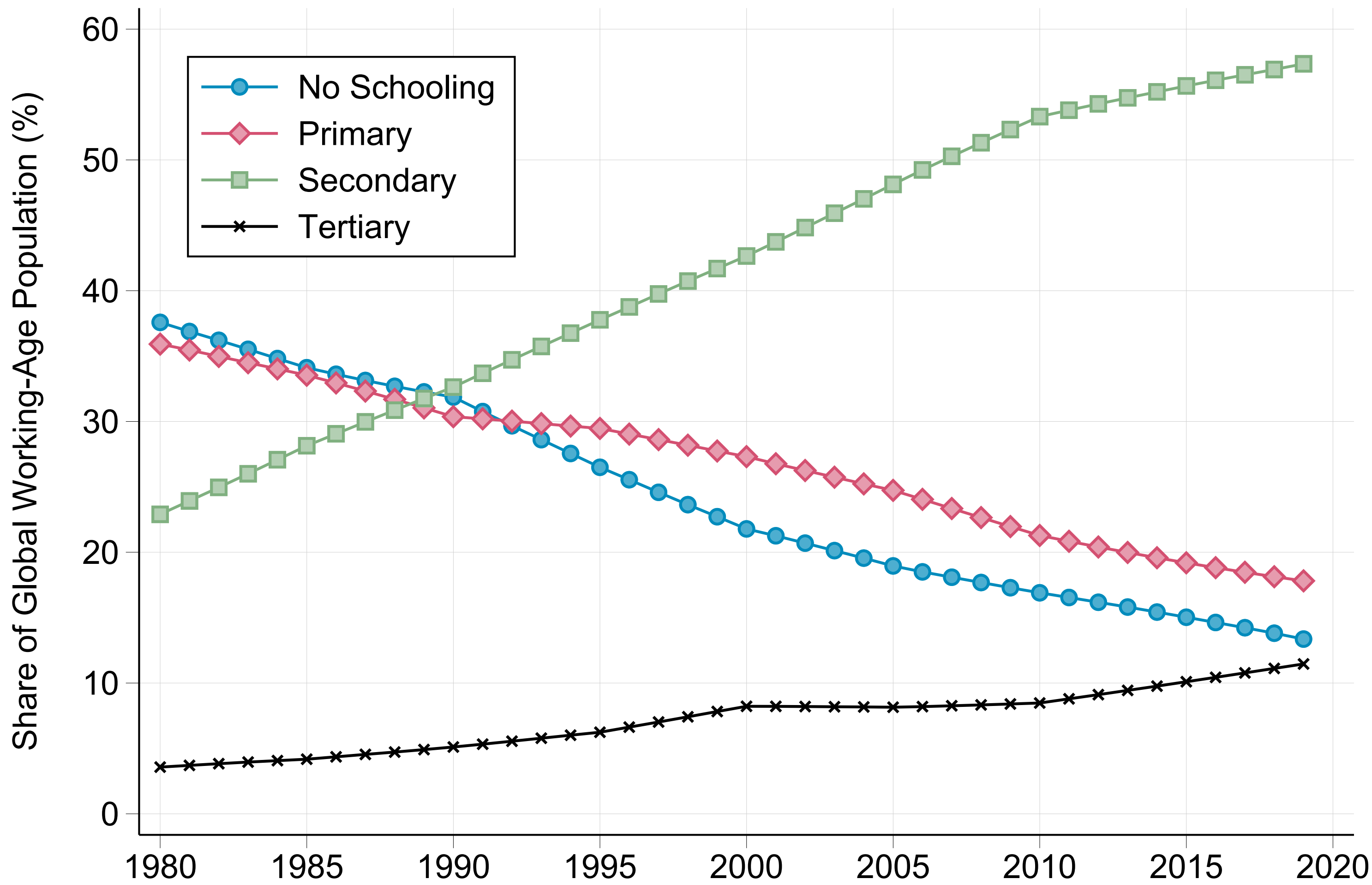
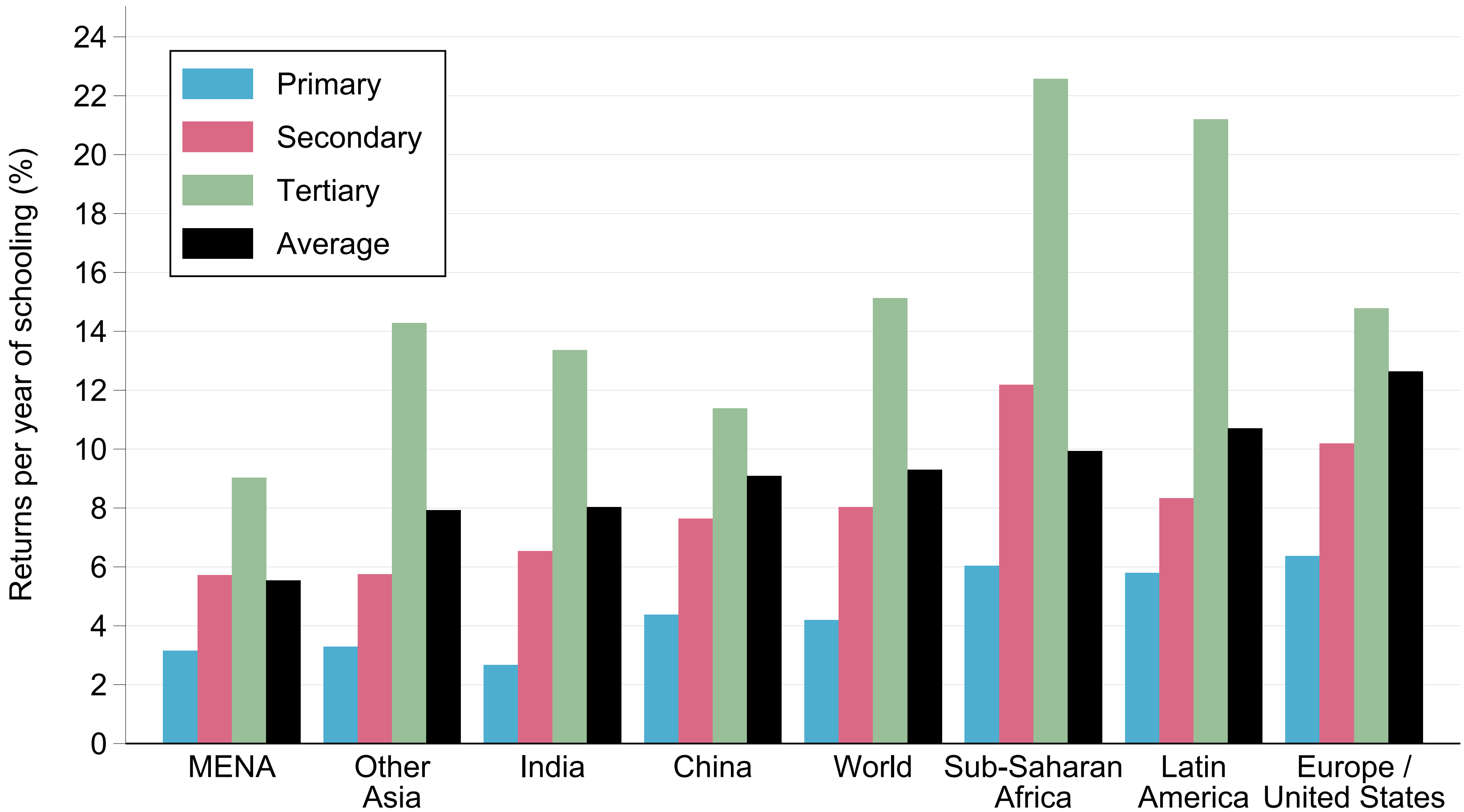


Figure 3 – Educational Attainment of the Global Working-Age Population, 1980-2019



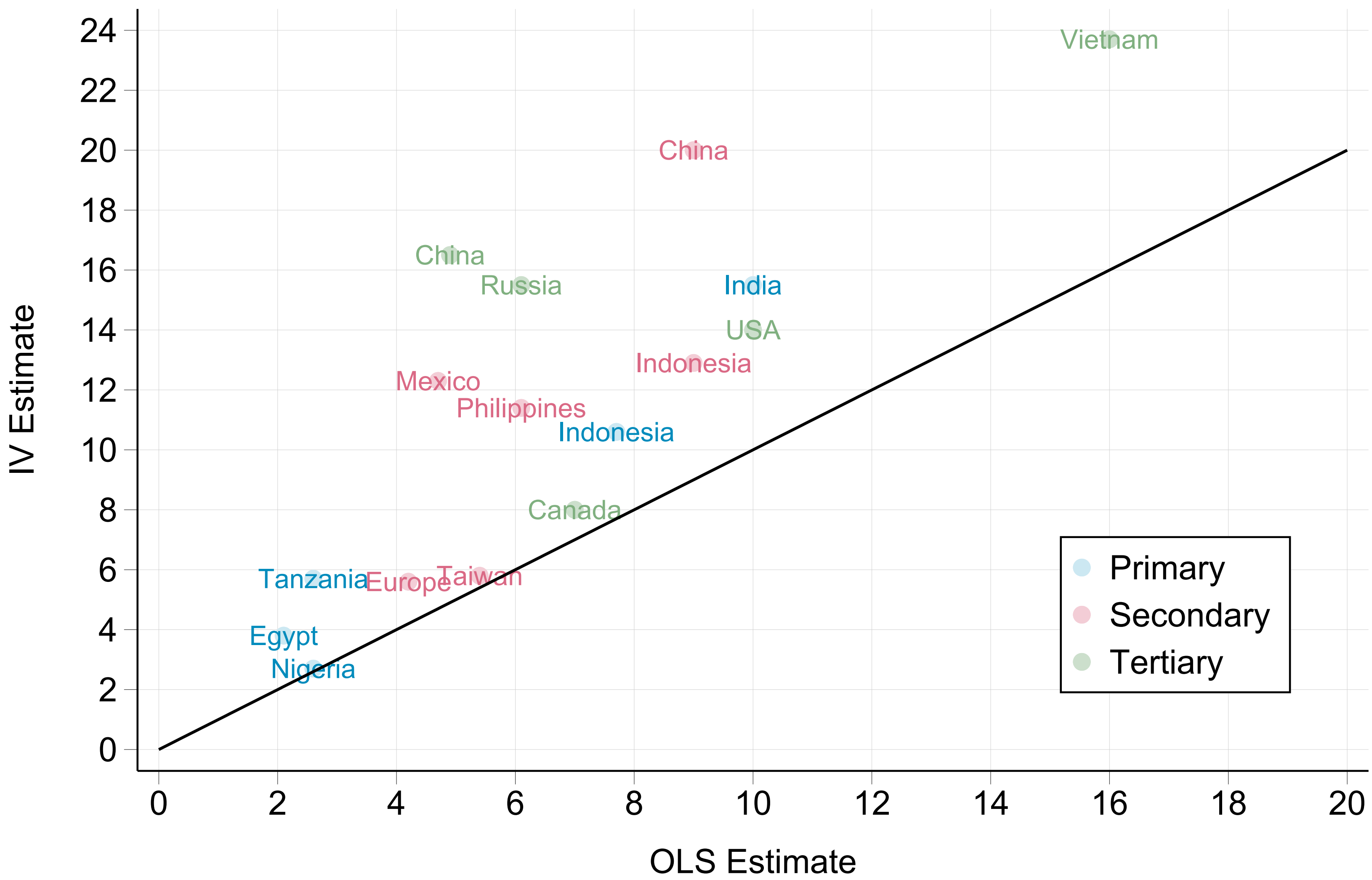
Notes. Author's computations combining data from Barro and Lee (2013) and updates, own sources, and working-age population estimates from the United Nations' World Population Prospects statistics.

Figure 4 – Returns to Schooling by World Region



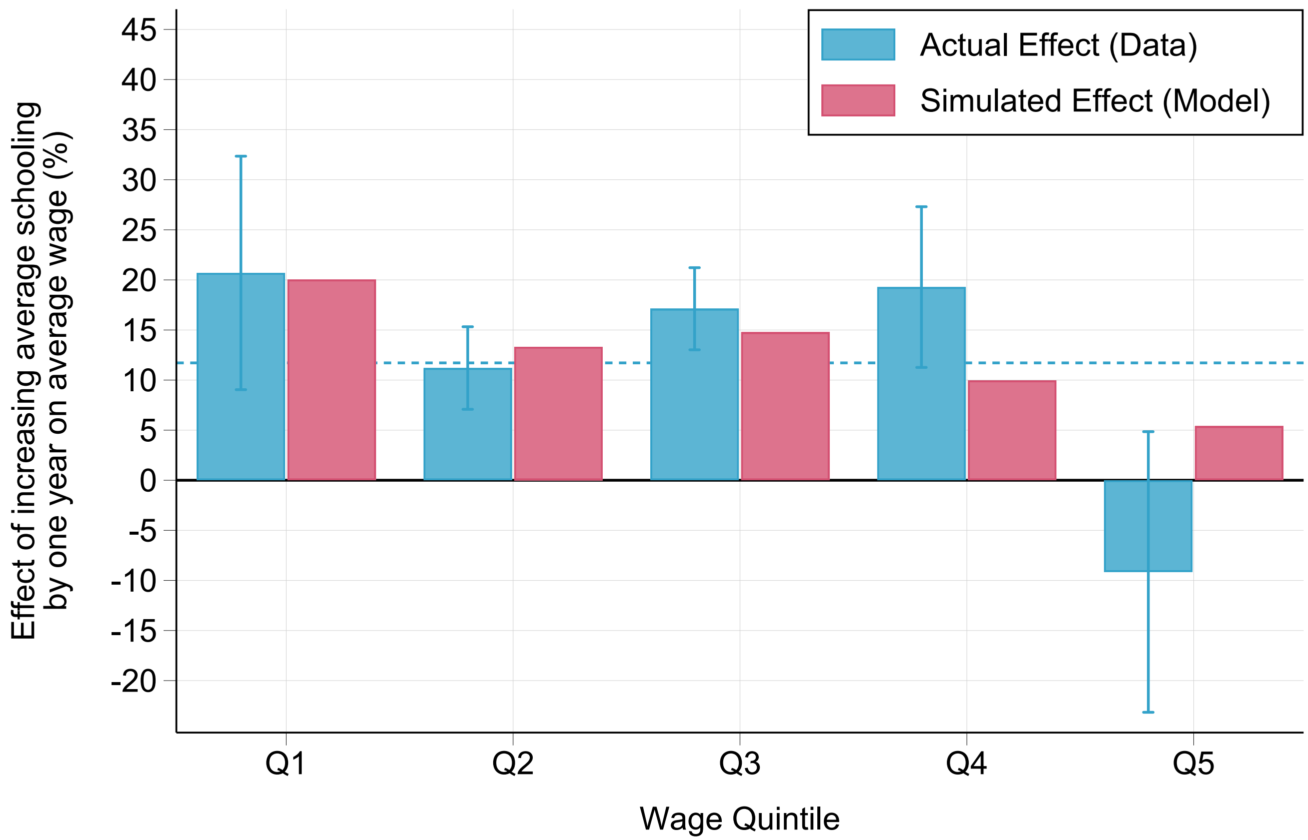
Notes. Author's computations using labor force survey microdata. Estimates correspond to the effect of one additional year of schooling on the log of personal income, estimated using modified Mincerian equations controlling for an experience quartic, gender, and interactions between the experience quartic and gender. Primary: return to a year of schooling among individuals with either no schooling, some primary education, or completed primary education. Secondary: return to a year of schooling among individuals with either some primary education, completed primary education, some lower or upper secondary education, or completed upper secondary education. Tertiary: return to a year of schooling among individuals with some upper secondary education, completed upper secondary education, some tertiary education, or completed tertiary education. Population-weighted averages of coefficients estimated in each country.

Figure 5 – Returns to Schooling: OLS versus IV Estimates



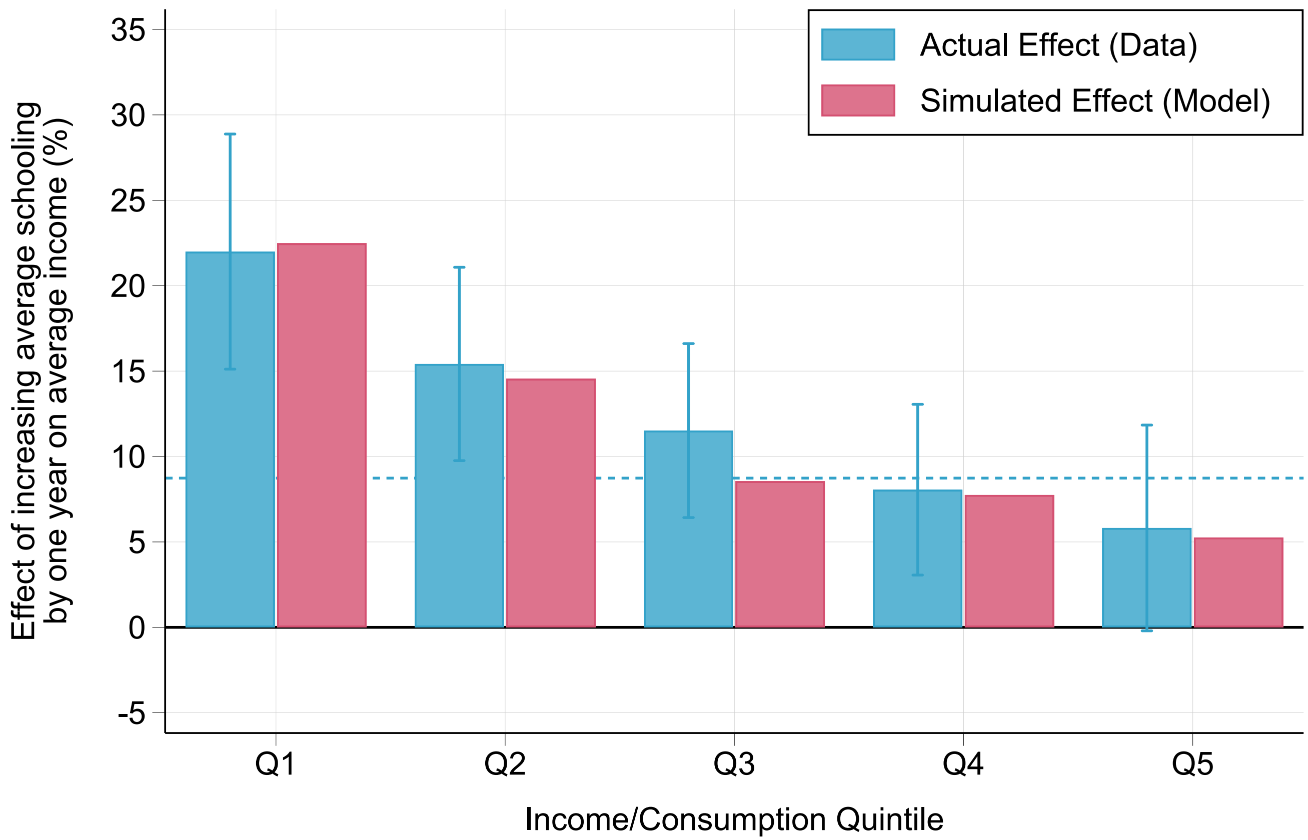
Notes. Author's elaboration compiling estimates from a number of empirical studies: see appendix table G4. The figure compares ordinary least squares (x-axis) and instrumental variable (y-axis) estimates of the return to an additional year of schooling. OLS estimates generally correspond to results from a Mincerian equation of the log of earnings on years of schooling, estimated over the entire working-age population. In contrast, IV estimates typically rely on quasi-experimental variation in access to a specific level of education (primary, secondary, or tertiary).

Figure 6 – Validation: Actual Versus Simulated Distributional Effects of India's District Primary Education Program



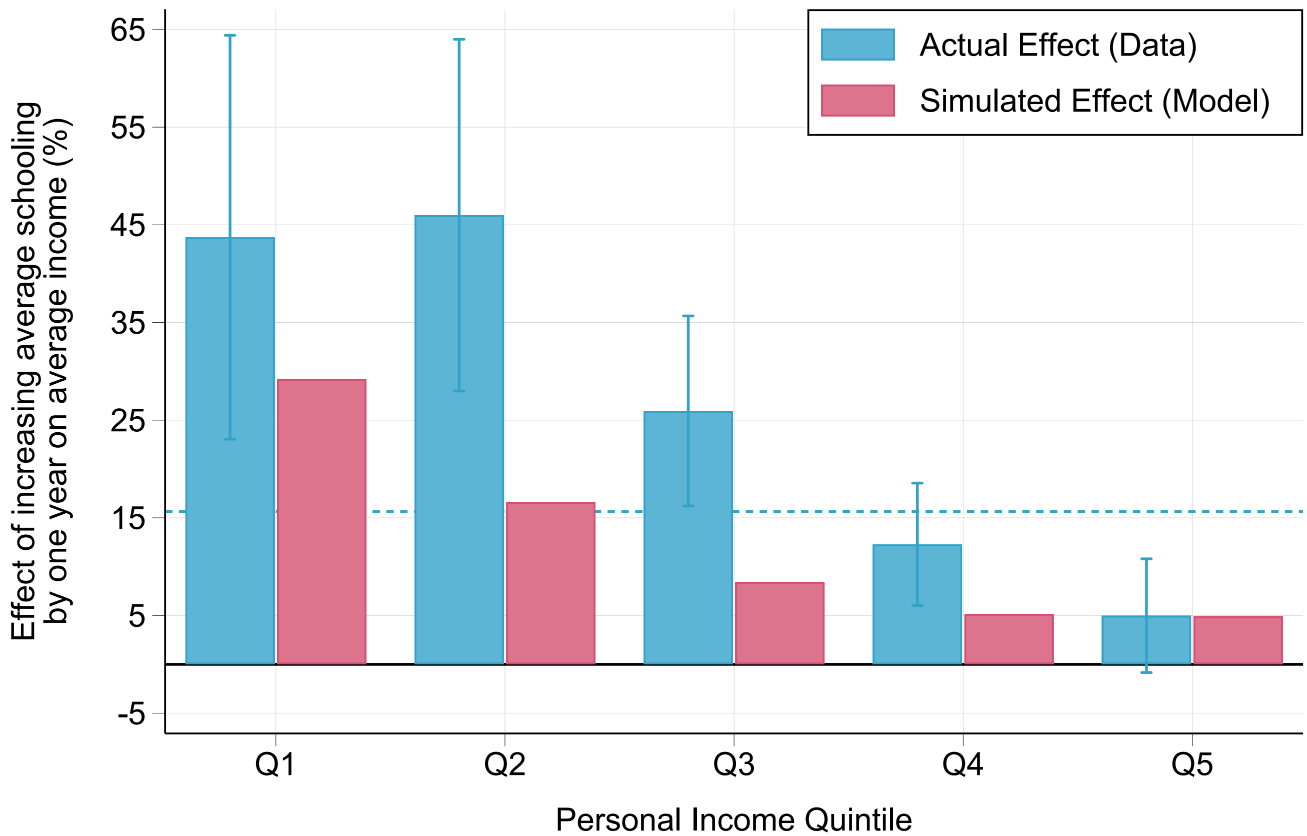
Notes. Actual effect: effect of increasing average district schooling by one year on individual wages by decile, instrumenting schooling with exposure to the District Primary Education Program. Capped spikes correspond to 95% confidence intervals. The dashed line shows the estimated effect of average years of schooling on average earnings. Estimates combine NES microdata with exposure to the policy from Khanna (2023). Simulated effect: predicted effect of increasing average schooling by one year (through primary education) on personal income by decile, calibrating the model on 2019 labor force survey microdata. Simulations assume returns to schooling of 13%.

Figure 7 – Validation: Actual Versus Simulated Distributional Effects of Indonesia's INPRES School Construction Program



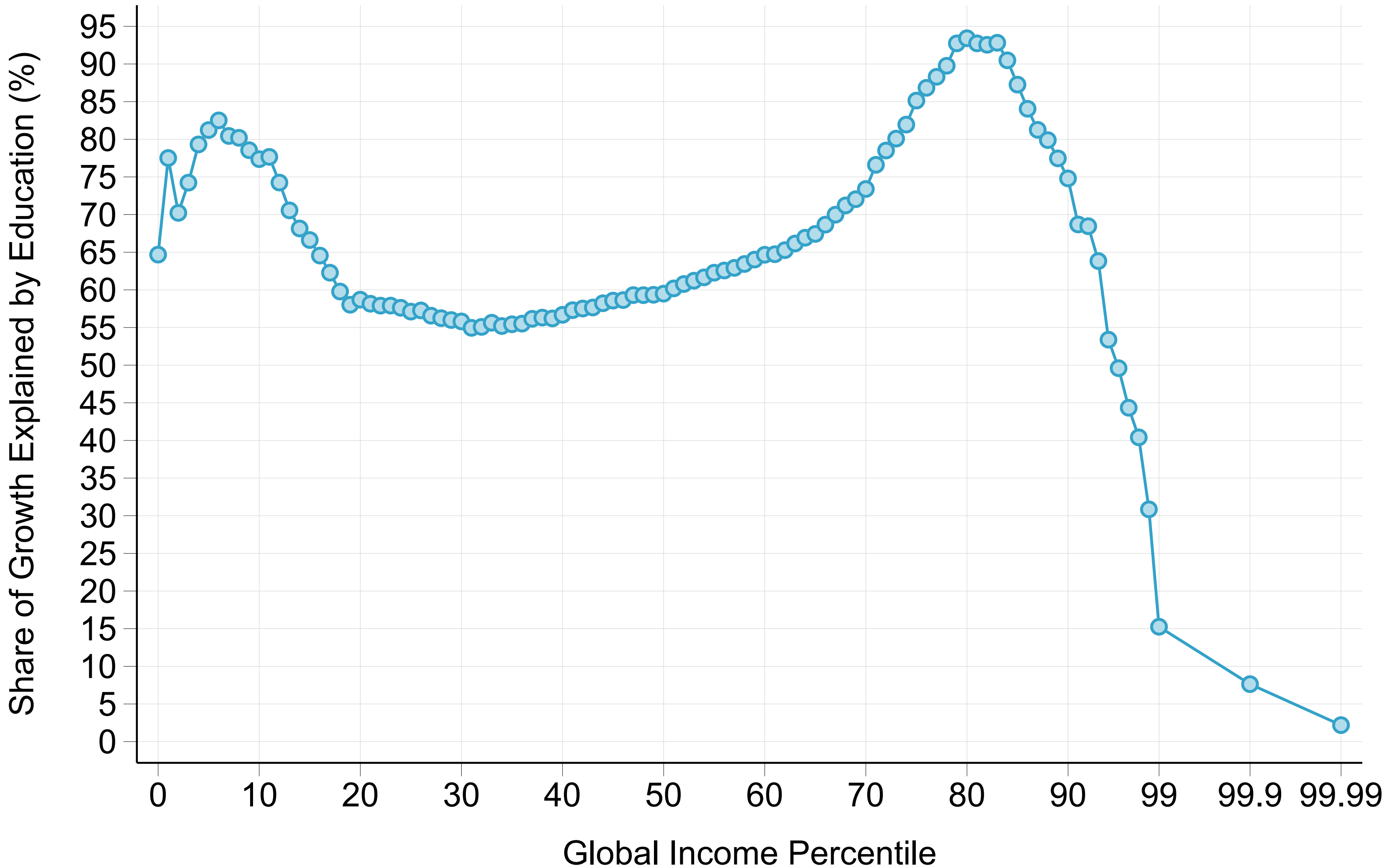
Notes. Actual effect: effect of increasing average district schooling by one year on per-adult consumption by decile, instrumenting schooling with exposure to the INPRES program. Capped spikes correspond to 95% confidence intervals. The dashed line shows the estimated effect of average years of schooling on average consumption. Estimates combine SUSENAS 1993-2019 microdata with INPRES program intensity from Duflo (2001). Simulated effect: predicted effect of increasing average schooling by one year (through primary education) on personal income by decile, calibrating the model on 1996 labor force survey microdata. Simulations assume returns to schooling of 11%.

Figure 8 – Validation: Actual Versus Simulated Distributional Effects of U.S. Compulsory Schooling Laws



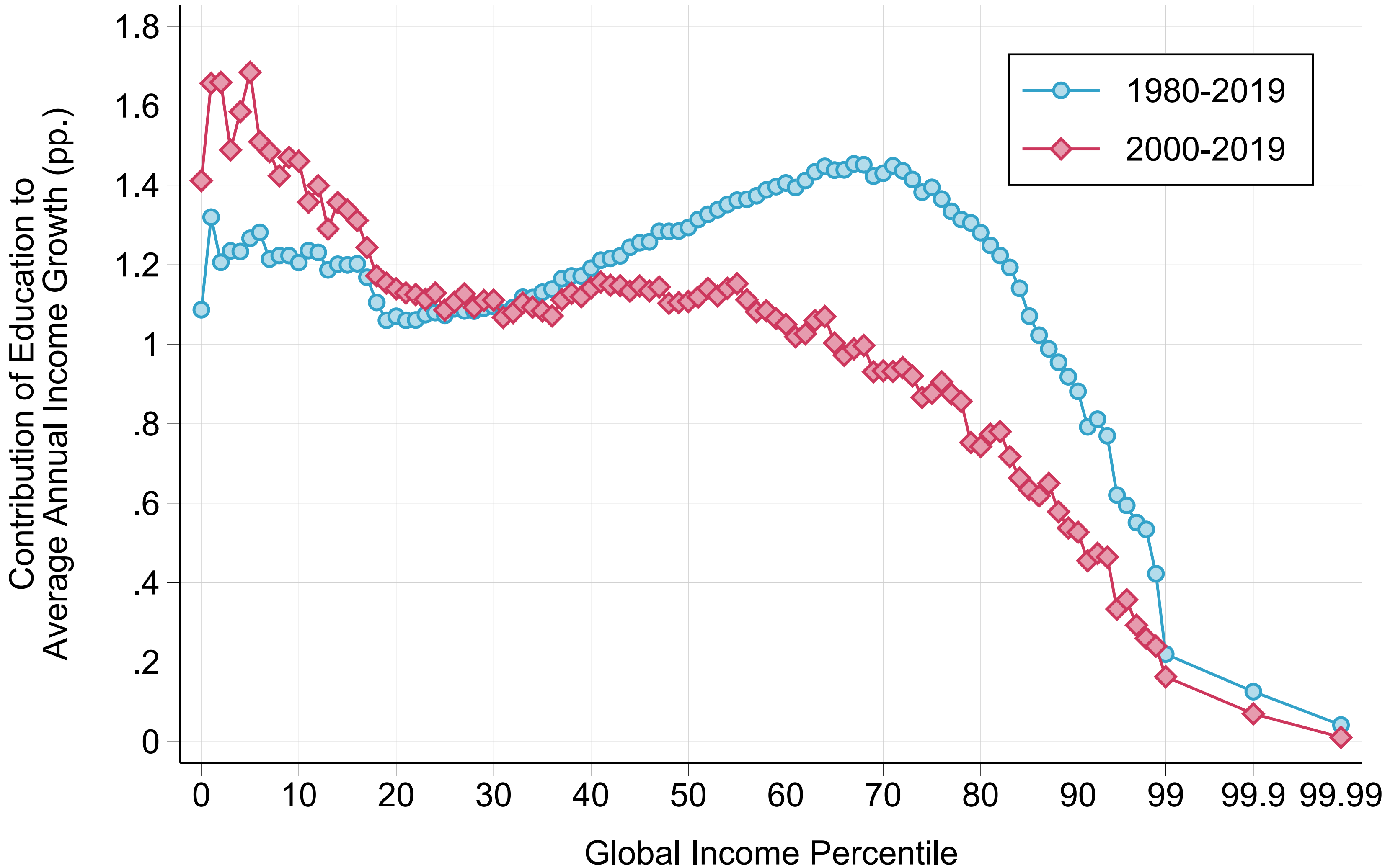
Notes. Actual effect: effect of increasing average state schooling by one year on personal income by decile, instrumenting schooling with state compulsory schooling laws. Capped spikes correspond to 95% confidence intervals. The dashed line shows the estimated effect of average years of schooling on average consumption. Estimates combine 1940-2000 census microdata with compulsory schooling laws from Acemoglu and Angrist (2000) and Clay, Lingwall, and Stephens (2021). Simulated effect: predicted effect of increasing average schooling by one year (through secondary education) on personal income by decile, calibrating the model on 1960 census microdata. Simulations assume returns to schooling of 12%.

Figure 9 – Share of Growth Explained by Education by Global Income Percentile, 1980-2019



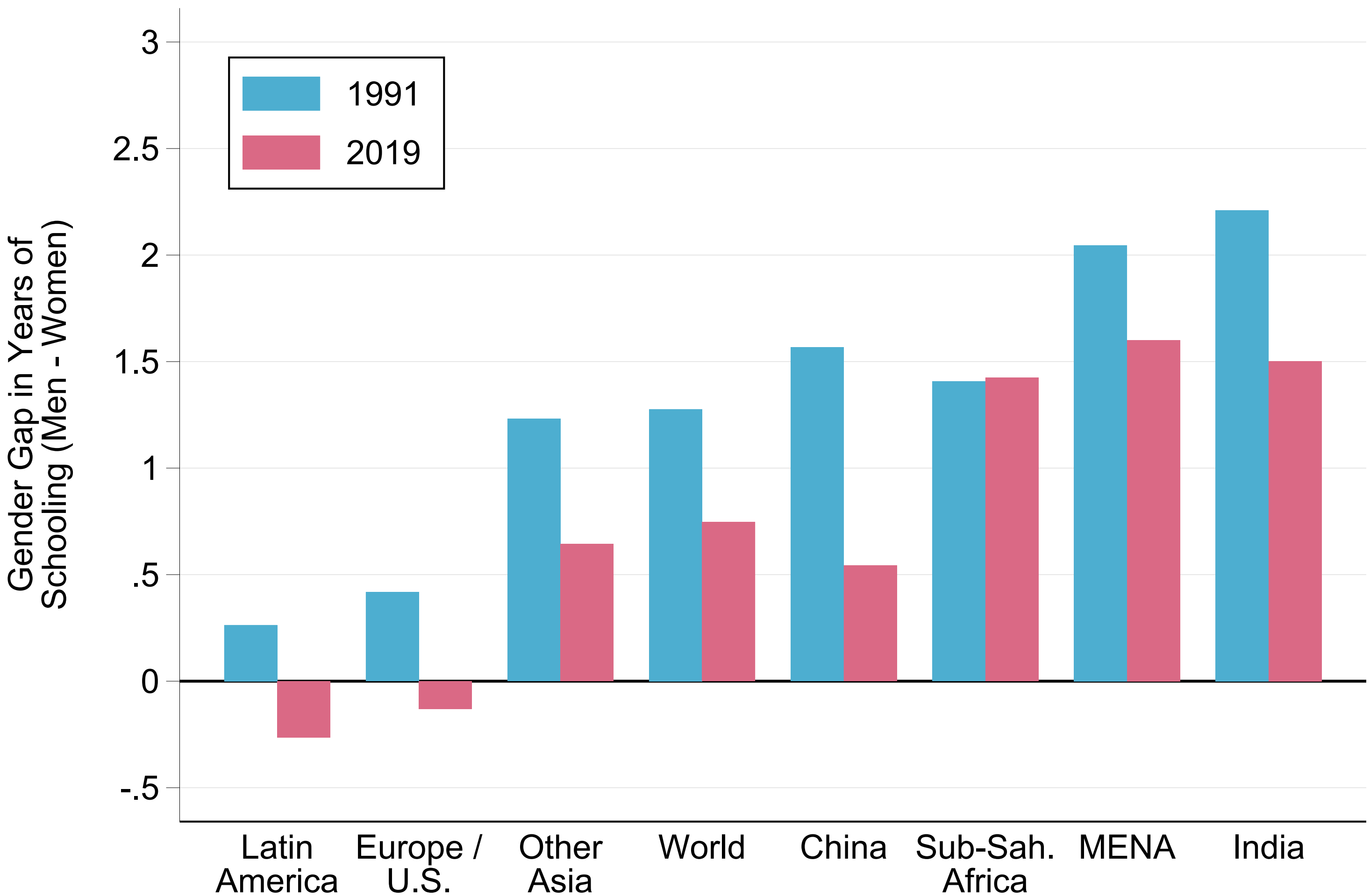
Notes. The figure plots the share of real pretax income growth that can be explained by improvements in educational attainment by global income percentile. Each point corresponds to the ratio of gains from schooling—equal to actual minus counterfactual income growth absent educational expansion—over actual income growth over the 1980-2019 period, calculated for each percentile of the world distribution of income.

Figure 10 – Contribution of Education to Global Average Annual Income Growth, 1980-2019



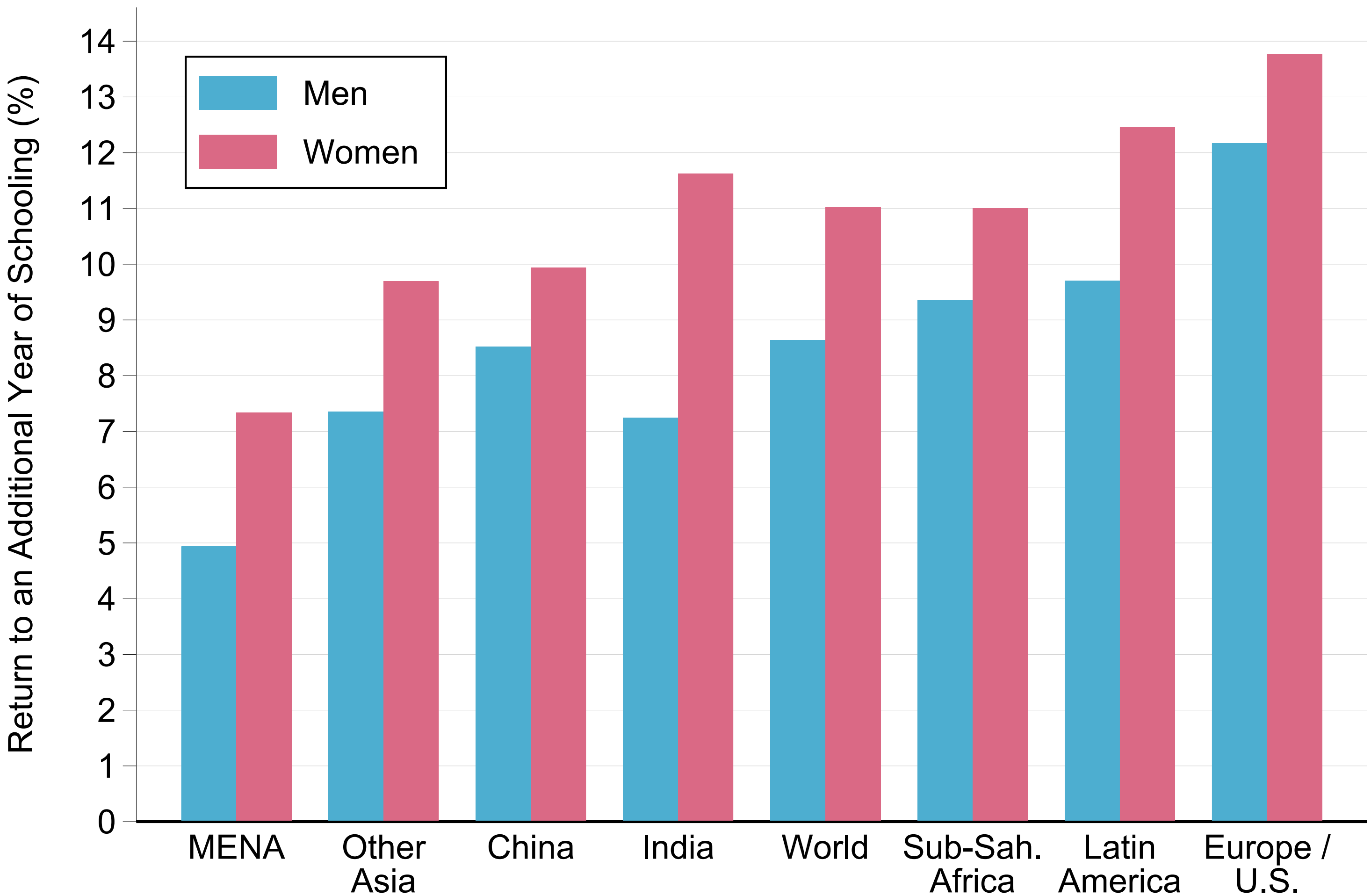
Notes. The figure plots annualized gains from schooling by global income percentile, calculated by taking the percent difference between actual and counterfactual income growth absent educational expansion, and annualizing the resulting figure over the period considered.

Figure 11 – Global Gender Schooling Inequality, 1991-2019
Gender Gap in Years of Schooling (Men - Women)



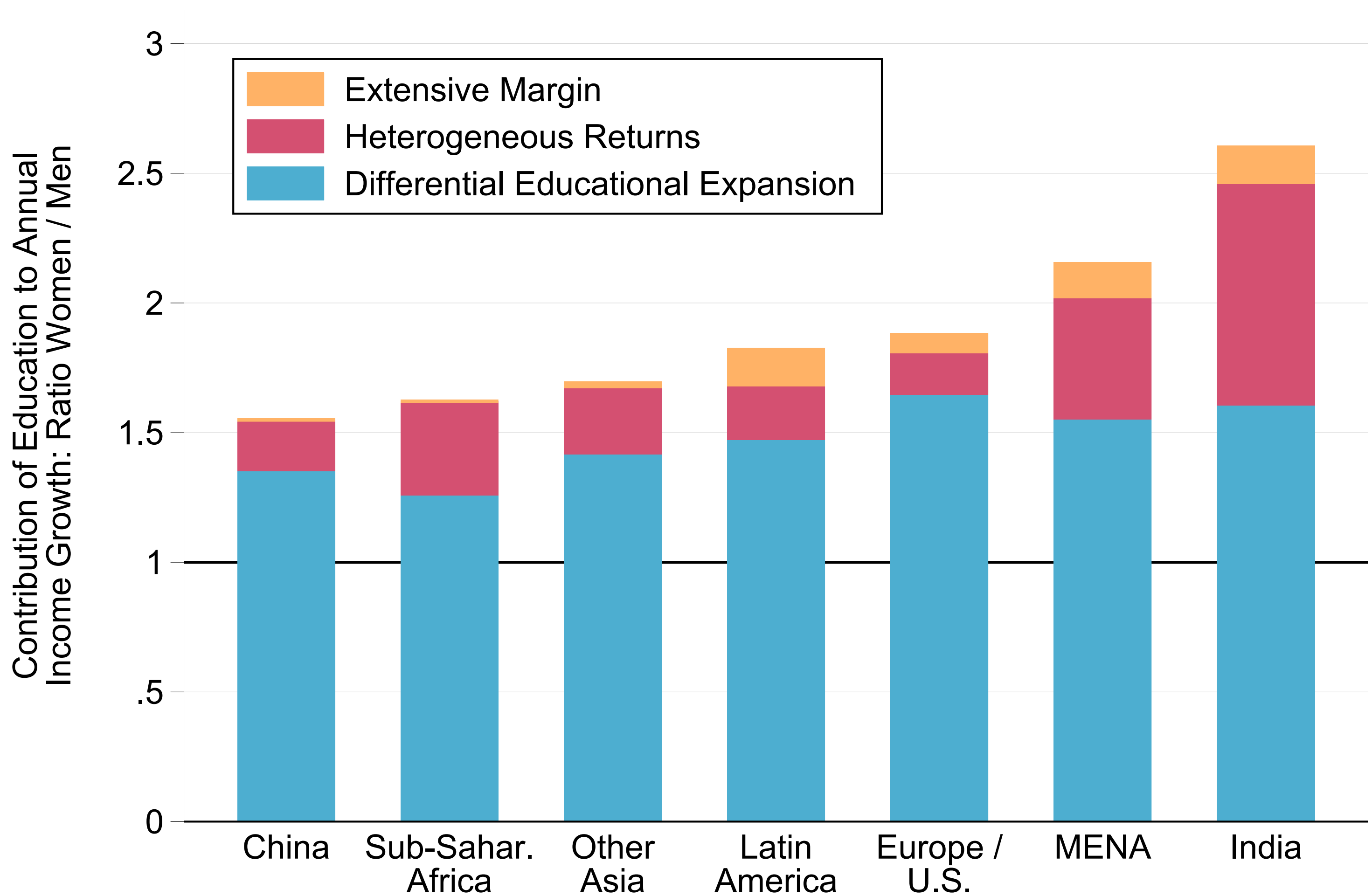
Notes. Author's computations using data from Barro and Lee (2013) and updates. The figure shows the population-weighted average gap in years of schooling between working-age men and women by world region and in the world as a whole.

Figure 12 – Returns to Schooling by Gender and World Region



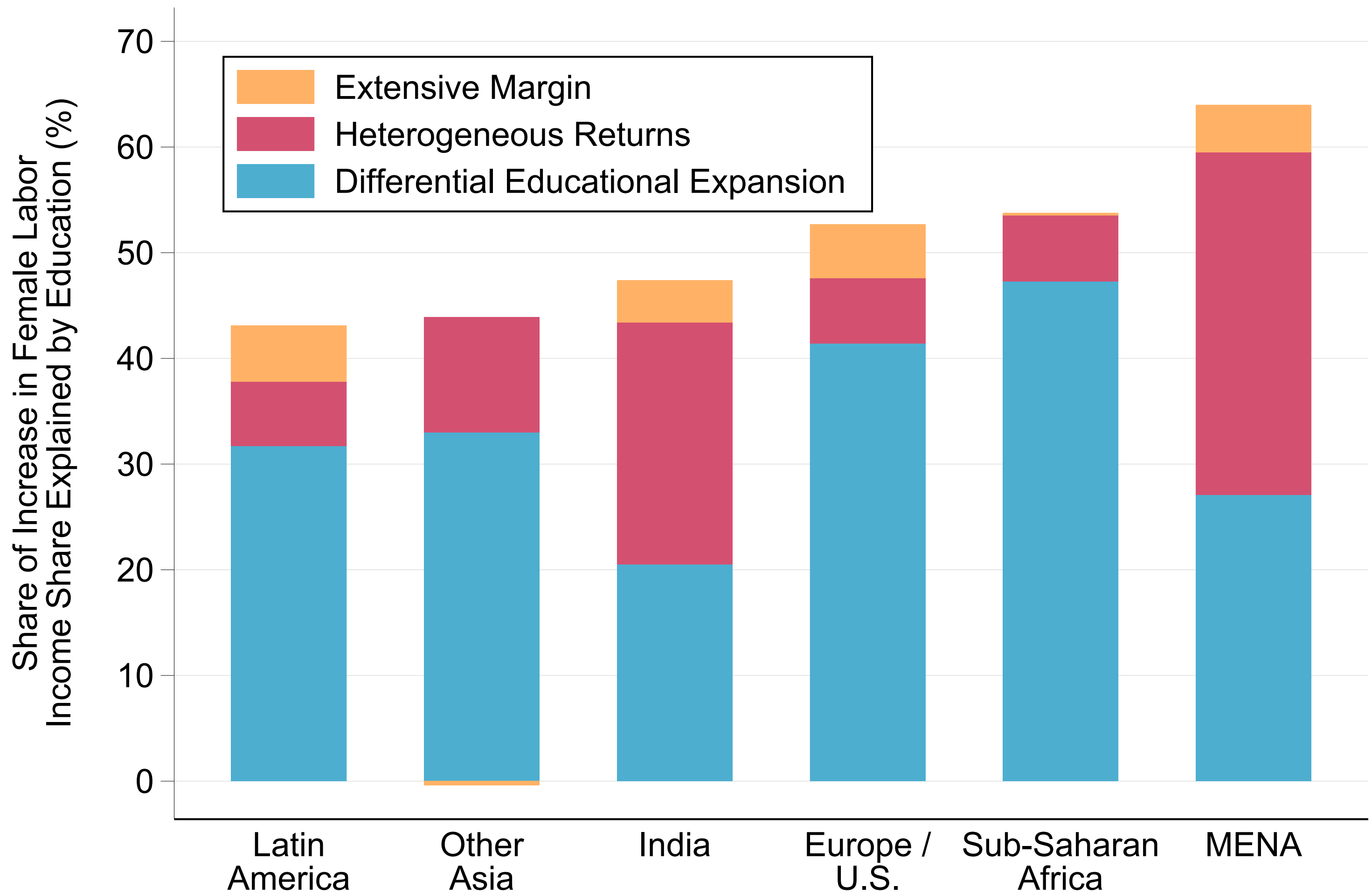
Notes. Author's computations using labor force survey microdata. Estimates correspond to the effect of one additional year of schooling on the log of personal income, estimated separately by gender using modified Mincerian equations controlling for an experience quartic. Population-weighted averages of coefficients estimated in each country.

Figure 13 – Contribution of Schooling to Gender Inequality Reduction by World Region
Gender Ratio of Income Gains from Schooling (Women / Men), 1991-2019



Notes. Author's computations using labor force survey microdata. Each bar corresponds to the ratio of gains from schooling among women over gains from schooling among men. Gains from schooling correspond to the percent difference between actual income and counterfactual income absent educational expansion. Population-weighted averages of gains from schooling estimated in each country.

Figure 14 – Share of Gender Inequality Reduction Explained by Education by World Region, 1991-2019



Notes. Author's computations using labor force survey microdata. Population-weighted averages of gains from schooling estimated in each country.

Table 1 – Survey Microdata Descriptive Statistics

	Countries	Observations	Share of Population Covered
Europe	39	743,328	100.0%
Northern America	2	539,862	100.0%
Latin America	24	4,126,194	96.5%
Asia-Pacific	29	2,524,531	95.5%
Middle East and North Africa	14	817,958	74.2%
Sub-Saharan Africa	42	876,054	98.9%
World	150	9,627,927	95.2%

Notes. The table reports the number of countries covered by the survey microdata, the total number of observations, and the share of the total population covered by world region and in the world as a whole (last row).

Table 2 – Distributional Growth Accounting, World, 1980-2019

	Total Income Growth (%)	Growth Without Education (%)	Contribution of Education (pp.)	Share of Growth Explained (%)
	g	$\tilde{g}$	$g-\tilde{g}$	$\frac{g-\tilde{g}}{g}$
Full Population	+98%	+45%	53	54%
Bottom 50%	+164%	+68%	96	59%
Bottom 20%	+115%	+36%	79	69%
Next 30%	+176%	+75%	101	57%
Middle 40%	+94%	+21%	73	78%
Top 10%	+91%	+59%	32	35%
Top 1%	+131%	+114%	17	13%
Top 0.1%	+173%	+161%	13	7%
Top 0.01%	+278%	+272%	6	2%

Notes. The table reports actual real income growth rates, counterfactual growth rates absent educational expansion, and the corresponding share of growth explained by education for different groups of the world distribution of income.

Table 3 – Education and Global Poverty Reduction

	1980	2019	Difference (%)	Share of Decline Explained (%)
Global Poverty: $2.15 / Day				
Actual	20%	9%	-55%	
Counterfactual	20%	13%	-32%	43%
Global Poverty: $3.65 / Day				
Actual	40%	14%	-65%	
Counterfactual	40%	21%	-46%	28%
Global Poverty: $6.85 / Day				
Actual	58%	27%	-53%	
Counterfactual	58%	40%	-30%	43%

Notes. The table compares the actual evolution of the global poverty headcount ratio to the evolution it would have followed absent educational expansion since 1980. All global poverty headcount ratios calculated using 2017 PPP USD. The income concept is pretax national income, as reported in the World Inequality Database. See appendix table A3 for comparable results using per-capita consumption distributions from the World Bank.

Table 4 – Distributional Growth Accounting by World Region and Country Income Group, 1980-2019

	Full Population			Bottom 20%		
	Actual Growth	Contrib. of Education	Share Explained	Actual Growth	Contrib. of Education	Share Explained
	g	$g-\tilde{g}$	$\frac{g-\tilde{g}}{g}$	g	$g-\tilde{g}$	$\frac{g-\tilde{g}}{g}$
Europe / Northern America	+81%	52	64%	+30%	82	>100%
Latin America	+38%	31	82%	+39%	60	>100%
China	+988%	312	32%	+377%	242	64%
India	+410%	106	26%	+208%	92	44%
Other Asia-Pacific	+117%	54	46%	+263%	113	43%
Middle East and North Africa	+118%	38	32%	+35%	38	>100%
Sub-Saharan Africa	+15%	30	>100%	+51%	57	>100%
Low-income	+16%	31	>100%	+50%	53	>100%
Low-middle-income	+182%	61	34%	+198%	96	48%
High-middle-income	+210%	85	40%	+218%	154	71%
High-income	+89%	52	59%	+77%	112	>100%

Notes. The table reports actual real income growth rates and the corresponding share of growth that can be explained by education, for the full population and the poorest 20%, by world region.

Table 5 – Education and Inequality Between and Within Countries

	1980	2019	Difference
Theil Index of Global Inequality			
Actual	1.06	1.08	0.02
Counterfactual	1.06	1.35	0.29
Between-Country Component			
Actual	0.60	0.34	-0.26
Counterfactual	0.60	0.33	-0.27
Within-Country Component			
Actual	0.46	0.74	0.28
Counterfactual	0.46	1.02	0.56
Within-Country Share (%)			
Actual	43%	69%	25
Counterfactual	43%	76%	32

Notes. The table compares the actual evolution of global inequality since 1980 to the evolution it would have followed absent educational expansion, decomposing these transformations into a between-country component and a within-country component. Within-country share: share of global inequality explained by inequality within countries.

Table 6 – From Standard to Distributional Growth Accounting

	Share of Growth Explained, 1980-2019	
	Global Average	Global Bottom 20%
Standard Growth Accounting Cross-Country Data, 10% Return	33%	23%
+ Adjusted Labor Share	43%	35%
+ Within-Country Inequality	43%	41%
+ Within-Country Labor Shares	43%	51%
+ Microdata, 2019 Returns	38%	38%
+ Distributional Effects, 2019 Returns	38%	52%
+ Distributional Effects, Adjusted Returns	53%	68%
+ Distributional Effects, IV Returns	55%	76%

Notes. The table reports the share of global economic growth and real income growth of the global bottom 20% that can be explained by education, depending on methodological assumptions and data sources used. Adjusted labor share: labor income includes mixed income. Within-country inequality: income distribution data from the World Inequality Database. Within-country labor shares: labor share varies by income group within each country. 2019 returns: Mincerian returns by level estimated using the microdata. Adjusted returns: true returns estimated by adding supply effects to 2019 returns. IV returns: 2019 returns adjusted using instrumental variable estimates of returns to schooling. Distributional effects: relative wage adjustments due to supply effects (imperfect substitution between skill groups).

Table 7 – Education and Global Gender Inequality, 1991-2019

	1991	2019	Diff.	Share Explained By Education	Share Explained (Cross-Country Average)
Global Female Labor Income Share	29.3%	32.1%	2.8		
Counterfactual: No Educational Progress	29.3%	30.6%	1.3	52%	46%
Counterfactual: + Heterogeneous Returns	29.3%	30.1%	0.8	73%	57%
Counterfactual: + Extensive Margin	29.3%	29.9%	0.6	80%	59%

Notes. The table reports actual versus counterfactual global female labor income shares under different assumptions. Global female labor income: total share of labor income received by women in the world as a whole. Change in education: only account for differential trends in schooling by gender, applying the same returns to schooling for men and women to build the counterfactual. Heterogeneous returns: account for differential returns by gender. Extensive margin: account for differential effects of schooling on employment by gender. Cross-country average: population-weighted average of the share of gender inequality reduction explained by education in each country.

Table 8 – Public Policies and Global Poverty Reduction:
Combining Direct Redistribution and Indirect Investment Benefits from Education

	1980	2019	Change (%)	Total Share of Change Explained (%)
Global Poverty Rate ($2.15/Day)				
Pretax Income Absent Educational Expansion	20%	13%	-32%	
Pretax Income	20%	8.7%	-55%	
Posttax Income	17%	5.1%	-70%	55%
Global Bottom 20% Average Income ($/Day)				
Pretax Income Absent Educational Expansion	1.3	1.8	+37%	
Pretax Income	1.3	2.8	+115%	
Posttax Income	1.5	4.0	+164%	78%
Global Bottom 50% Average Income ($/Day)				
Pretax Income Absent Educational Expansion	2.7	4.6	+69%	
Pretax Income	2.7	7.1	+163%	
Posttax Income	2.9	8.8	+200%	66%

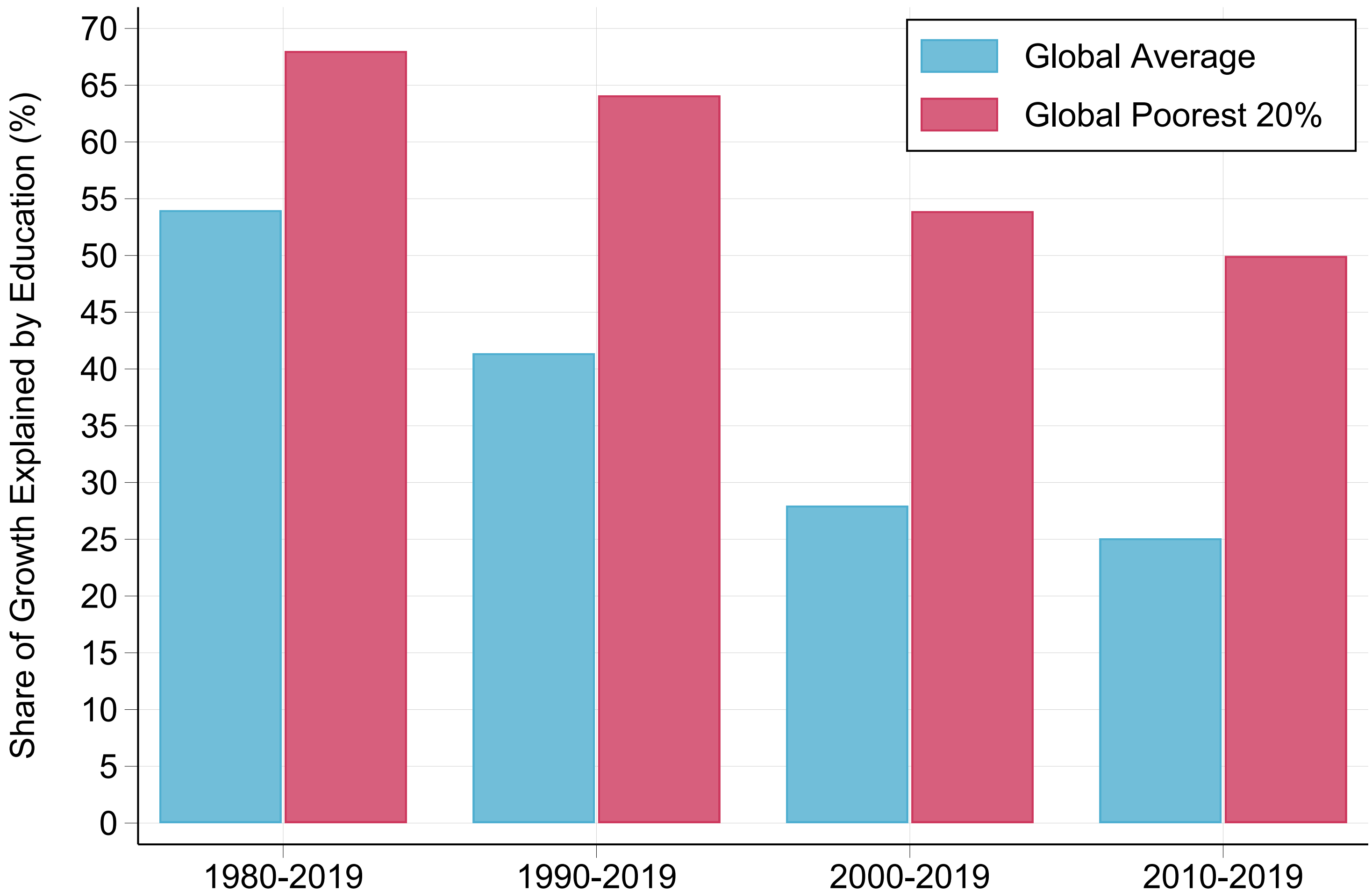
Notes. The table compares the evolution of global poverty and the average income of the global bottom 20% and bottom 50% under three scenarios. The first one considers the evolution of each indicator if there had been no educational progress since 1980 ("pretax income absent educational expansion"). The second one corresponds to the actual evolution of each indicator in terms of pretax income ("pretax income"). The third one corresponds to the actual evolution of each indicator in terms of posttax income, that is, after removing all taxes and adding all cash and in-kind transfers (see Gethin, 2023). The last column displays the corresponding share of global poverty reduction or real income gains that can be attributed to public policies, combining indirect investment benefits from education (moving from "pretax income absent educational expansion" to pretax income) and direct redistribution (moving from pretax to posttax income), calculated as one minus the ratio of the first row to the third row of the fourth column within each panel. Global poverty rate calculated at $2.15 per day in 2017 PPP USD. Real incomes of the bottom 20% and bottom 50% expressed in 2021 PPP USD as in the rest of the paper. Estimates of the world distribution of income from the World Inequality Database. See appendix table A12 for comparable results using per-capita consumption distributions from the World Bank.

Appendices

A. Additional Figures and Tables

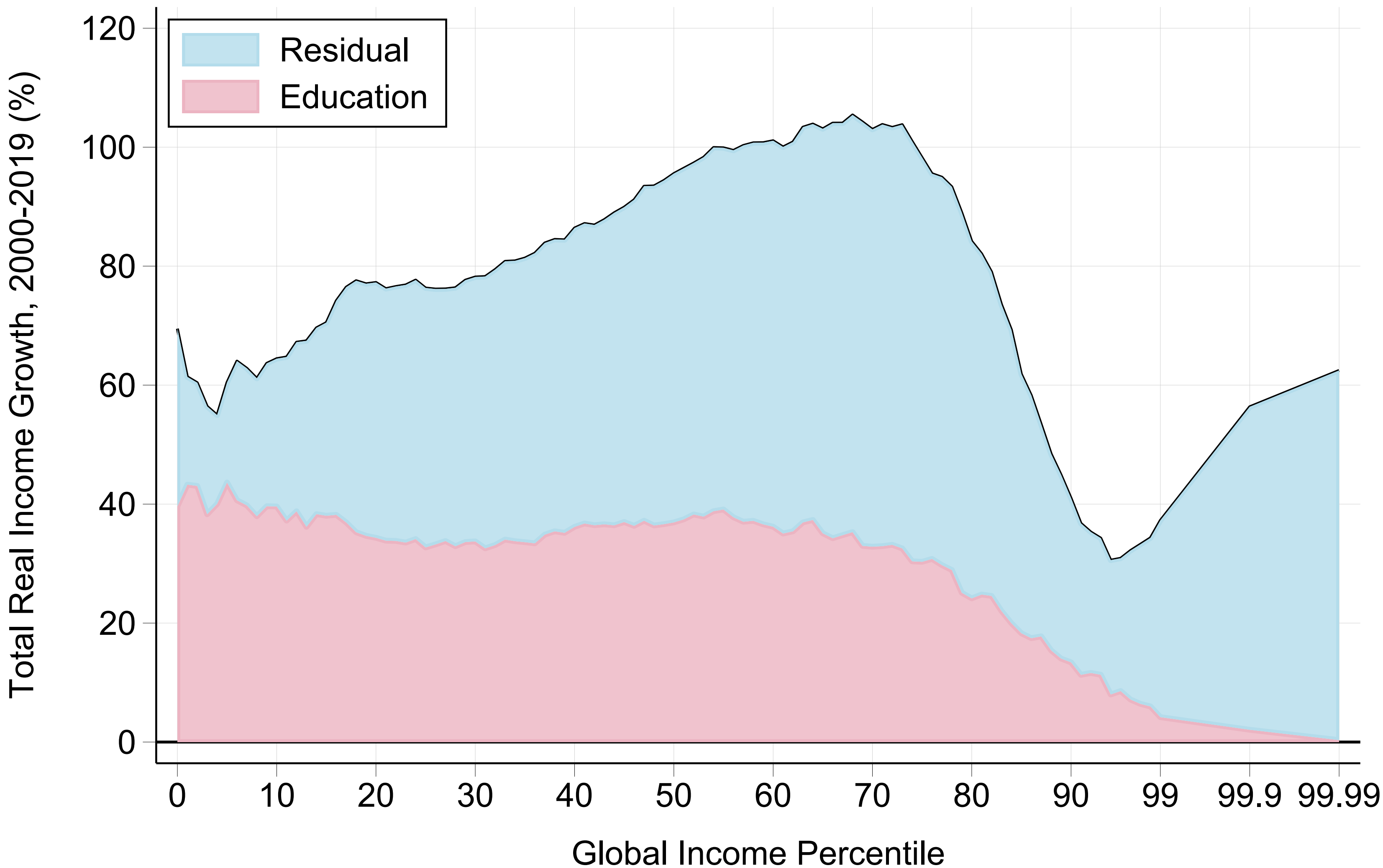
A.1. Additional Results: World Distribution of Income

Figure A1 – Growth Accounting, 1980-2019: Global Average vs. Poorest 20%



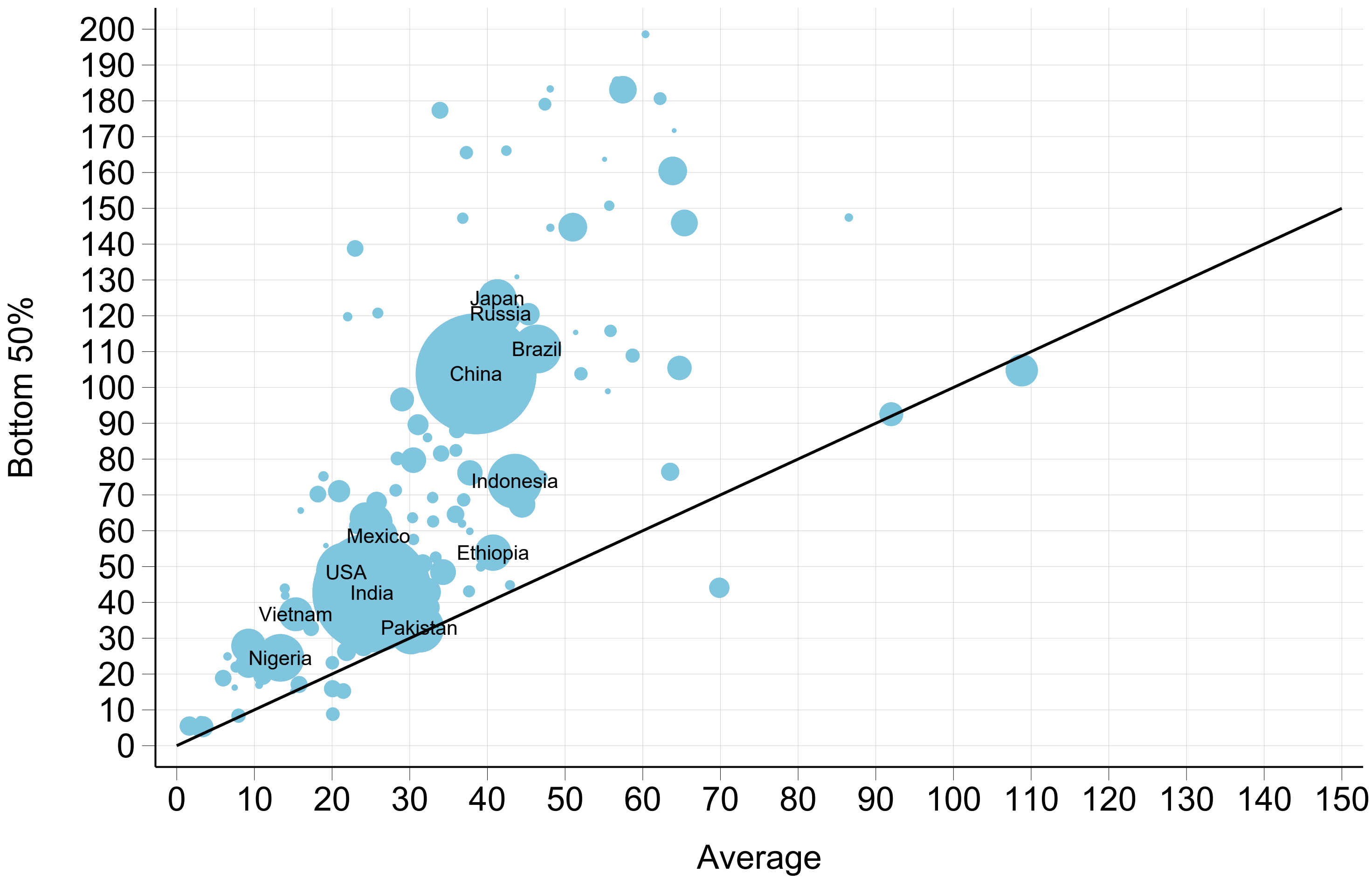
Notes. Author's calculations.

Figure A2 – The Distribution of Global Economic Growth, 2000-2019



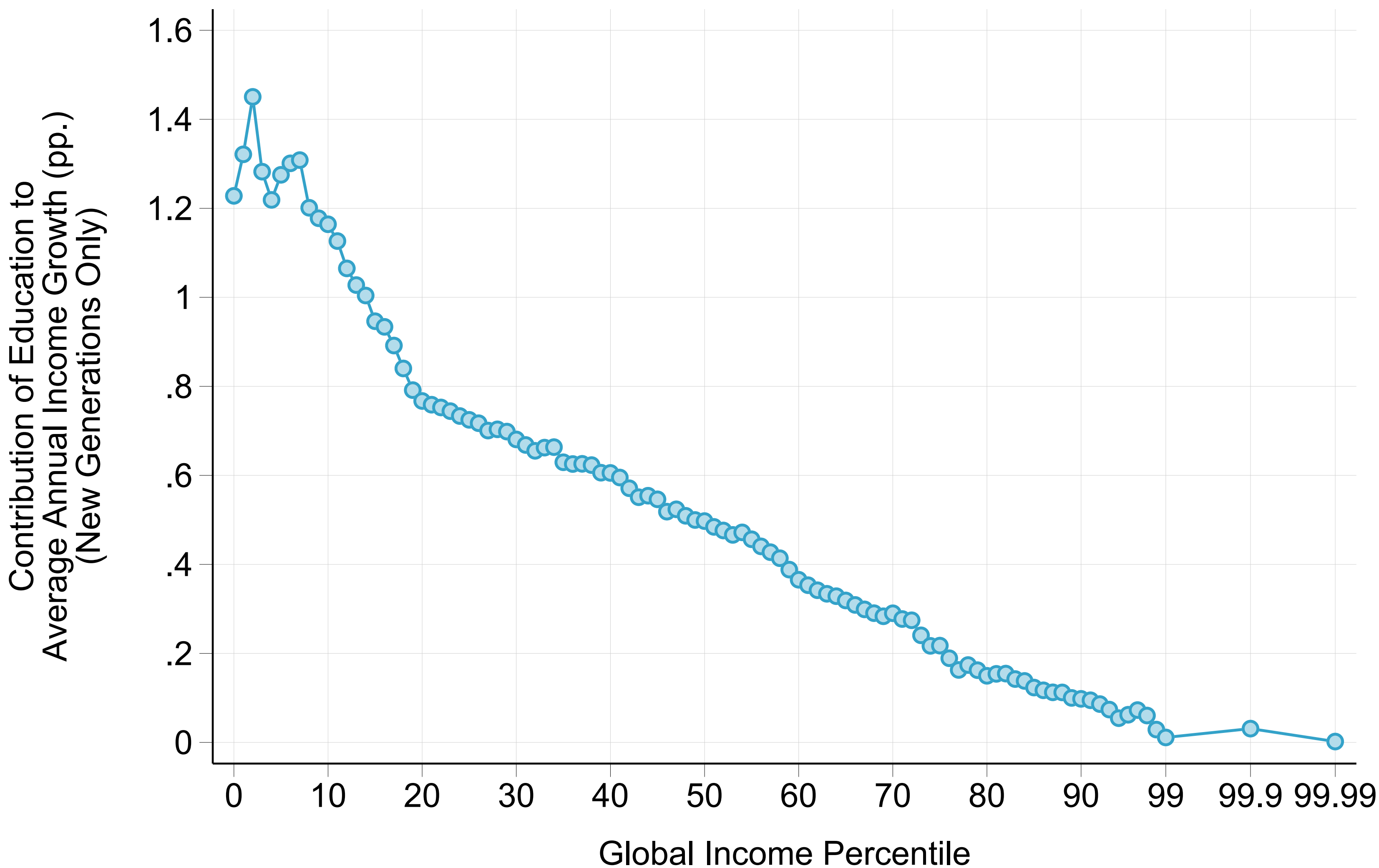
Notes. The figure shows total real income growth by global income percentile, and decomposes it into a part that can be explained by education and an unexplained component. The income concept is pretax national income.

Figure A3 – Income Gains From Schooling by Country, 1980-2019: Average Versus Bottom 50%



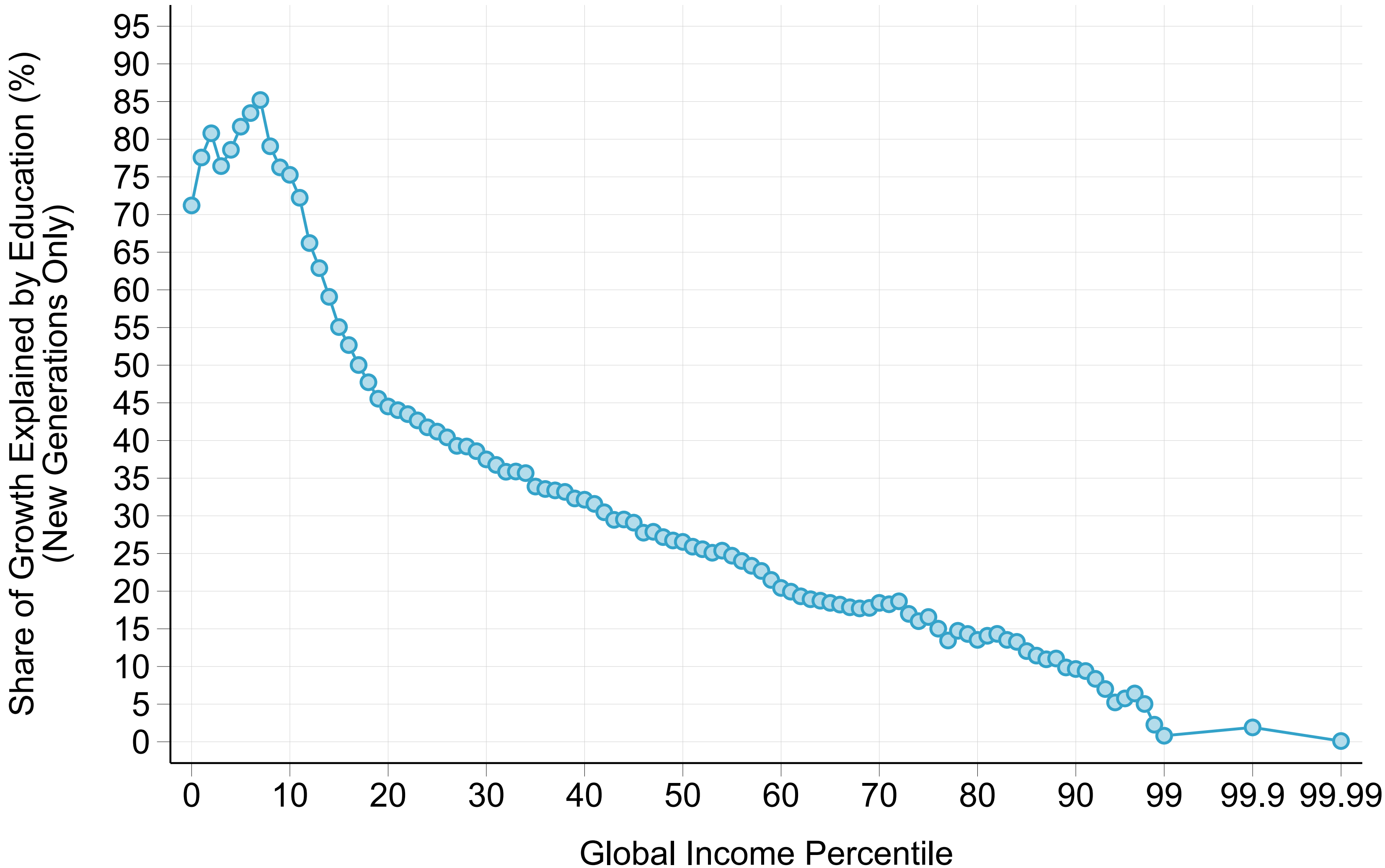
Notes. Author's calculations. The figure compares gains from schooling for the population as a whole (average) versus the bottom 50% in each country over the 1980-2019 period. Gains from schooling correspond to the percent increase in income generated by educational expansion since 1980.

Figure A4 – Contribution of Education to Global Average Annual Income Growth, 1980-2019:
Educational Progress Among Post-1980 Generations Only



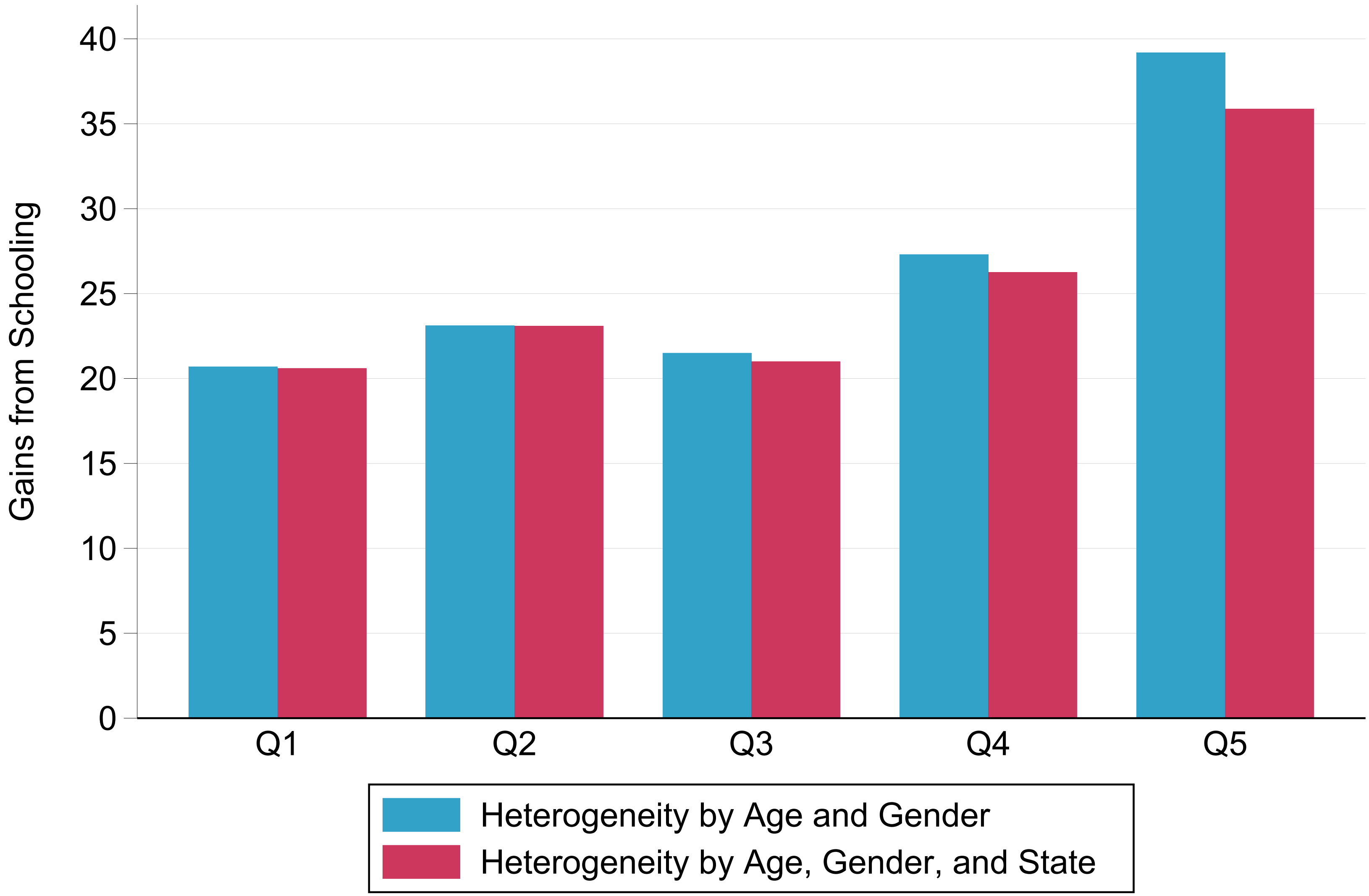
Notes. Author's calculations. The figure reports schooling gains by global income percentile, focusing on educational expansion among post-1980 generations only.

Figure A5 – Share of Growth Explained by Education by Global Income Percentile, 1980-2019: Educational Progress Among Post-1980 Generations Only



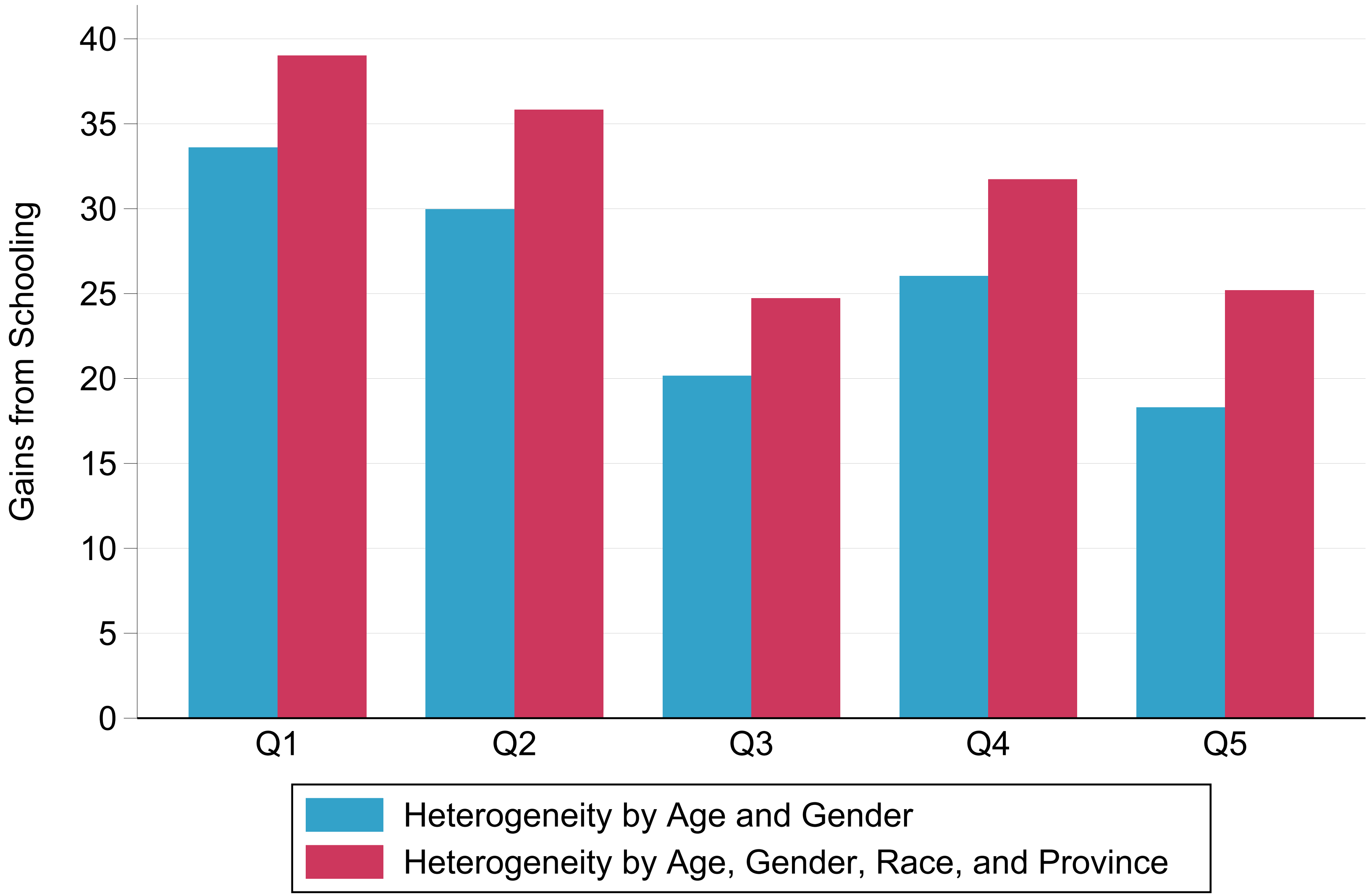
Notes. Author's calculations. The figure reports the share of growth explained by education by global income percentile, focusing on educational expansion among post-1980 generations only.

Figure A6 – Gains from Schooling With and Without Heterogeneous Educational Expansion by Socioeconomic Characteristic: India, 1983-2019



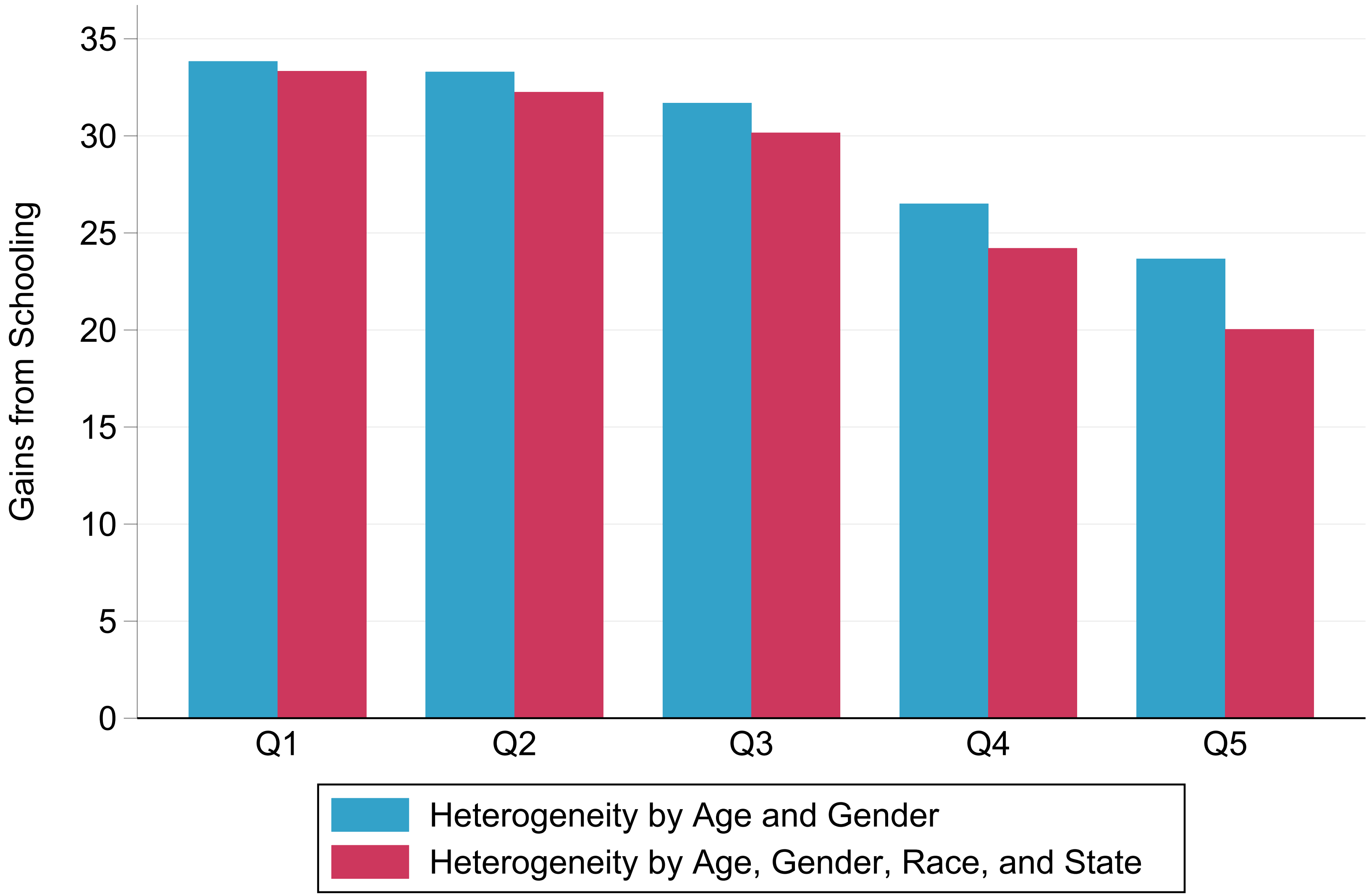
Notes. Educational attainment by age, gender, and state of residence in 1983 is estimated using the 1983 National Sample Survey. Education levels of individuals in the 2019 Periodic Labor Force Survey are then downgraded by age-gender cell (specification 1) or age-gender-state cell (specification 2) until reaching 1983 levels. Their earnings are reduced using estimates of returns to schooling by level. Finally, the figure plots schooling gains by income quintile, defined as the percent difference between actual income and counterfactual income absent educational expansion since 1983.

Figure A7 – Gains from Schooling With and Without Heterogeneous Educational Expansion by Socioeconomic Characteristic: South Africa, 2002-2019



Notes. Educational attainment by age, gender, race, and province of residence in 2002 is estimated using the 2002 General Household Survey. Education levels of individuals in the 2019 General Household Survey are then downgraded by age-gender cell (specification 1) or age-gender-race-province cell (specification 2) until reaching 2002 levels. Their earnings are reduced using estimates of returns to schooling by level. Finally, the figure plots schooling gains by income quintile, defined as the percent difference between actual income and counterfactual income absent educational expansion since 2002.

Figure A8 – Gains from Schooling With and Without Heterogeneous Educational Expansion by Socioeconomic Characteristic: United States; 1980-2019



Notes. Educational attainment by age, gender, race, and state of residence in 1980 is estimated using 1980 IPUMS census sample microdata. Education levels of individuals in the 2019 Current Population Survey are then downgraded by age-gender cell (specification 1) or age-gender-race-state cell (specification 2) until reaching 1980 levels. Their earnings are reduced using estimates of returns to schooling by level. Finally, the figure plots schooling gains by income quintile, defined as the percent difference between actual income and counterfactual income absent educational expansion since 1980.

Table A1 – Distributional Growth Accounting, World: 2000-2019

	Total Income Growth (%)	Growth Without Education (%)	Contribution of Education (pp.)	Share of Growth Explained (%)
	g	$\tilde{g}$	$g - \tilde{g}$	$\frac{g-\tilde{g}}{g}$
Full Population	+55%	+39%	15	28%
Bottom 50%	+82%	+46%	36	44%
Bottom 20%	+69%	+31%	38	55%
Next 30%	+84%	+49%	35	42%
Middle 40%	+78%	+53%	25	32%
Top 10%	+37%	+30%	7	19%
Top 1%	+42%	+39%	4	8%
Top 0.1%	+57%	+55%	2	3%
Top 0.01%	+62%	+62%	0.3	0.5%

Notes. The table reports actual real income growth rates, counterfactual growth rates absent educational expansion, and the corresponding share of growth explained by education for different groups of the world distribution of income.

Table A2 – Distributional Growth Accounting, World: World Bank Data

	1980-2019			2000-2019		
	Total Income Growth (%)	Growth Without Education (%)	Share of Growth Explained (%)	Total Income Growth (%)	Growth Without Education (%)	Share of Growth Explained (%)
	g	$\tilde{g}$	$1-\frac{\tilde{g}}{g}$	g	$\tilde{g}$	$1-\frac{\tilde{g}}{g}$
Full Population	+77%	+18%	76%	+48%	+29%	39%
Bottom 50%	+205%	+79%	61%	+97%	+54%	45%
Bottom 20%	+196%	+77%	61%	+83%	+39%	53%
Next 30%	+207%	+80%	61%	+102%	+58%	43%
Middle 40%	+83%	+7%	92%	+77%	+47%	39%
Top 10%	+62%	+20%	68%	+27%	+16%	39%
Top 1%	+77%	+43%	44%	+26%	+19%	26%
Top 0.1%	+105%	+64%	39%	+19%	+11%	41%
Top 0.01%	+154%	+118%	24%	+5%	+3%	50%

Notes. The table reports actual real income growth rates, counterfactual growth rates absent educational expansion, and the corresponding share of growth explained by education for different groups of the world distribution of income. Estimates of the world distribution of income constructed from World Bank per-capita income and consumption data.

Table A3 – Education and Global Poverty Reduction: World Bank Data

	1980	2019	Difference (%)	Share of Decline Explained (%)
Global Poverty: $2.15 / Day				
Actual	46%	10%	-79%	
Counterfactual	46%	24%	-48%	39%
Global Poverty: $3.65 / Day				
Actual	60%	25%	-59%	
Counterfactual	60%	45%	-24%	58%
Global Poverty: $6.85 / Day				
Actual	70%	47%	-33%	
Counterfactual	70%	64%	-8%	75%

Notes. The table compares the actual evolution of the global poverty headcount ratio to the evolution it would have followed absent educational expansion since 1980. All global poverty headcount ratios calculated using 2017 PPP USD. Estimates of the world distribution of income constructed from World Bank per-capita income and consumption data.

Table A4 – Distributional Growth Accounting, World: Alternative Elasticities of Substitution

	1980-2019			2000-2019		
	Low Substitutability	Benchmark	High Substitutability	Low Substitutability	Benchmark	High Substitutability
Full Population	58%	53%	51%	31%	28%	27%
Bottom 50%	82%	57%	51%	66%	43%	38%
Bottom 20%	95%	68%	60%	81%	54%	48%
Next 30%	80%	56%	50%	63%	41%	37%
Middle 40%	99%	76%	71%	43%	32%	31%
Top 10%	20%	34%	35%	7%	19%	19%
Top 1%	0%	12%	12%	0%	8%	8%
Top 0.1%	3%	7%	7%	0%	3%	3%
Top 0.01%	0.5%	2%	2%	0%	0.5%	0.5%

Notes. The table reports the share of growth explained by education for different groups of the world distribution of income, depending on assumptions made on the substitutability of skilled and unskilled workers. Low substitutability: $\sigma_1 = 1.5$, $\sigma_2 = 2.5$, $\sigma_3 = 4$. Benchmark: $\sigma_1 = \sigma_2 = \sigma_3 = 5$. High substitutability: $\sigma_1 = 5$, $\sigma_2 = 7$, $\sigma_3 = 9$.

Table A5 – Distributional Growth Accounting, World, 1980-2019: Alternative Nesting of CES Production Function

	Total Income Growth (%)	Growth Without Education (%)	Contribution of Education (pp.)	Share of Growth Explained (%)
	g	$\tilde{g}$	$g - \tilde{g}$	$\frac{g-\tilde{g}}{g}$
Full Population	+98%	+46%	51	53%
Bottom 50%	+164%	+85%	79	48%
Bottom 20%	+115%	+49%	66	57%
Next 30%	+176%	+93%	82	47%
Middle 40%	+94%	+28%	66	70%
Top 10%	+91%	+55%	36	40%
Top 1%	+131%	+108%	23	18%
Top 0.1%	+173%	+158%	16	9%
Top 0.01%	+278%	+271%	6	2%

Notes. The table reports actual real income growth rates, counterfactual growth rates absent educational expansion, and the corresponding share of growth explained by education for different groups of the world distribution of income.

Table A6 – Distributional Growth Accounting, World, 1980-2019:
With Capital Income Affected by Education

	Total Income Growth (%)	Growth Without Education (%)	Contribution of Education (pp.)	Share of Growth Explained (%)
	g	$\tilde{g}$	$g - \tilde{g}$	$\frac{g-\tilde{g}}{g}$
Full Population	+98%	+32%	66%	67%
Bottom 50%	+164%	+55%	109%	66%
Bottom 20%	+115%	+25%	90%	78%
Next 30%	+176%	+62%	114%	65%
Middle 40%	+94%	+12%	82%	87%
Top 10%	+91%	+44%	47%	51%
Top 1%	+131%	+86%	45%	34%
Top 0.1%	+173%	+117%	57%	33%
Top 0.01%	+278%	+204%	73%	26%

Notes. The table reports actual real income growth rates, counterfactual growth rates absent educational expansion, and the corresponding share of growth explained by education for different groups of the world distribution of income. Returns to schooling are assumed to affect both labor income and capital income by the same amount.

Table A7 – Distributional Growth Accounting, World, 1980-2019: Educational Progress Among Post-1980 Generations Only

	Total Income Growth (%)	Growth Without Education (%)	Contribution of Education (pp.)	Share of Growth Explained (%)
	g	$\tilde{g}$	$g - \tilde{g}$	$\frac{g-\tilde{g}}{g}$
Full Population	+98%	+86%	12	12%
Bottom 50%	+164%	+103%	61	37%
Bottom 20%	+115%	+45%	70	61%
Next 30%	+176%	+117%	59	33%
Middle 40%	+94%	+79%	15	16%
Top 10%	+91%	+87%	4	4%
Top 1%	+131%	+130%	1	1%
Top 0.1%	+173%	+170%	3	2%
Top 0.01%	+278%	+277%	0.2	0.1%

Notes. The table reports actual real income growth rates, counterfactual growth rates absent educational expansion, and the corresponding share of growth explained by education for different groups of the world distribution of income. Educational expansion is defined as improvements in educational attainment among post-1980 generations only.

Table A8 – Cross-Country Correlates of Schooling: Cross-Sectional Estimates

	Expected Years of Schooling	Primary School Enrollment	Secondary School Enrollment
Log Public Education Expenditure Per Child	1.351*	10.916***	4.115**
	(0.699)	(1.391)	(1.949)
Government Effectiveness	0.554*	1.961	-2.093
	(0.317)	(1.638)	(2.288)
Trade-to-GDP Ratio	-0.547***	-0.624	-3.220***
	(0.150)	(0.785)	(1.136)
Internet Usage	0.292	-1.899	1.969
	(0.372)	(1.961)	(2.778)
Mobile cellular subscriptions (per 100 people)	-0.045	3.584**	1.987
	(0.294)	(1.535)	(2.152)
Skill Bias of Technology	-0.124	-1.207	-0.151
	(0.196)	(1.025)	(1.425)
Log GDP Per Capita	0.600	-4.813**	7.720**
	(0.539)	(2.368)	(3.392)
Child Population (% Total)	-0.820**	-1.457	-13.819***
	(0.347)	(1.699)	(2.426)
Constant	12.062***	84.302***	54.134***
	(0.510)	(2.632)	(3.714)
Treament			
Binary			
N	139	140	135
Adj. R-squared	0.75	0.54	0.80

Notes. All variables are standardized to have a mean of 0 and standard deviation of 1. Log public education expenditure per child: data from Gethin (2023). Skill bias of technology: average of relative efficiency terms A_H/A_L for primary, secondary, and tertiary education, weighted by the share of workers in each category, estimated using labor force survey microdata. Log GDP per capita: data from the World Inequality Database. Education expenditure and GDP expressed in 2019 PPP USD. All other variables: data from the World Bank Development Indicators. Standard errors in parenthesis. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A9 – Cross-Country Correlates of Schooling: Panel Estimates

	Expected Years of Schooling	Primary School Enrollment	Secondary School Enrollment
Log Public Education Expenditure Per Child	1.490***	6.110***	8.032***
	(0.246)	(1.321)	(1.527)
Government Effectiveness	-0.170*	0.328	-0.734
	(0.097)	(0.625)	(0.669)
Trade-to-GDP Ratio	0.157**	0.398	-0.446
	(0.061)	(0.378)	(0.353)
Internet Usage	-0.542***	-5.183***	-4.089***
	(0.053)	(0.319)	(0.370)
Mobile cellular subscriptions (per 100 people)	0.146***	0.228	1.369***
	(0.052)	(0.307)	(0.327)
Log GDP Per Capita	0.137	-2.125	2.472
	(0.241)	(1.381)	(1.639)
Child Population (% Total)	-0.527***	3.996***	-8.557***
	(0.147)	(0.910)	(1.054)
Constant	12.617***	91.593***	64.970***
	(0.110)	(0.506)	(0.801)
Treament			
Binary			
N	1,888	2,060	1,561
Adj. R-squared	0.95	0.86	0.97

Notes. All variables are standardized to have a mean of 0 and standard deviation of 1. Log public education expenditure per child: data from Gethin (2023). Log GDP per capita: data from the World Inequality Database. Education expenditure and GDP expressed in 2019 PPP USD. All other variables: data from the World Bank Development Indicators. Standard errors in parenthesis. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A10 – Income Gains from Schooling With and Without Relative Efficiency Gains, 2000-2019

	Average			Bottom 50%		
	Without Efficiency Gains	With Efficiency Gains	Ratio	Without Efficiency Gains	With Efficiency Gains	Ratio
Europe	+9%	+17%	1.88	+13%	+41%	3.12
United States	+4%	+4%	1.06	+6%	+8%	1.29
Brazil	+24%	+23%	0.93	+49%	+50%	1.03
Mexico	+10%	+8%	0.87	+20%	+18%	0.92
Other Latin America	+9%	+8%	0.96	+15%	+21%	1.43
Indonesia	+15%	+27%	1.86	+31%	+45%	1.44
Thailand	+16%	+28%	1.77	+27%	+42%	1.57
Ghana	+4%	+3%	0.74	+7%	+5%	0.77
South Africa	+13%	+9%	0.70	+37%	+38%	1.04
Average	+9%	+13%	1.38	+16%	+30%	1.88

Notes. The table compares income gains from schooling with and without relative efficiency gains in selected countries and groups of countries. With efficiency gains (backward growth accounting): income gains from schooling estimated by reducing incomes in 2019 to match education levels observed in 2000 (holding the relative skill bias to its 2019 level). Without efficiency gains (forward growth accounting): income gains from schooling estimated by increasing incomes in 2000 to match education levels observed in 2019 (holding the relative skill bias to its 2000 level). Europe: Austria, Belgium, Denmark, Estonia, Finland, France, Iceland, Ireland, Luxembourg, Norway, Portugal, Sweden, Switzerland, United Kingdom. Other Latin America: Argentina, Bolivia, Chile, Colombia, Costa Rica, Dominican Republic, Ecuador, Guatemala, Panama, Paraguay, Peru, Uruguay.

Table A11 – Share of Growth Explained by Education Without and With Relative Efficiency Gains, 2000-2019

	Average			Bottom 50%		
	Without Efficiency Gains	With Efficiency Gains	Ratio	Without Efficiency Gains	With Efficiency Gains	Ratio
Europe	51%	73%	1.45	60%	89%	1.49
United States	18%	19%	1.06	44%	57%	1.28
Brazil	91%	86%	0.95	>100%	>100%	1.00
Mexico	>100%	>100%	1.00	>100%	>100%	1.00
Other Latin America	22%	21%	0.96	28%	39%	1.35
Indonesia	21%	35%	1.67	52%	68%	1.30
Thailand	31%	49%	1.60	36%	50%	1.40
Ghana	8%	6%	0.75	15%	12%	0.79
South Africa	58%	41%	0.72	>100%	>100%	1.00
Average	36%	45%	1.25	46%	63%	1.37

Notes. The table compares the share of growth explain by education with and without relative efficiency gains in selected countries and groups of countries. With efficiency gains (backward growth accounting): income gains from schooling estimated by reducing incomes in 2019 to match education levels observed in 2000 (holding the relative skill bias to its 2019 level). Without efficiency gains (forward growth accounting): income gains from schooling estimated by increasing incomes in 2000 to match education levels observed in 2019 (holding the relative skill bias to its 2000 level). Europe: Austria, Belgium, Denmark, Estonia, Finland, France, Iceland, Ireland, Luxembourg, Norway, Portugal, Sweden, Switzerland, United Kingdom. Other Latin America: Argentina, Bolivia, Chile, Colombia, Costa Rica, Dominican Republic, Ecuador, Guatemala, Panama, Paraguay, Peru, Uruguay.

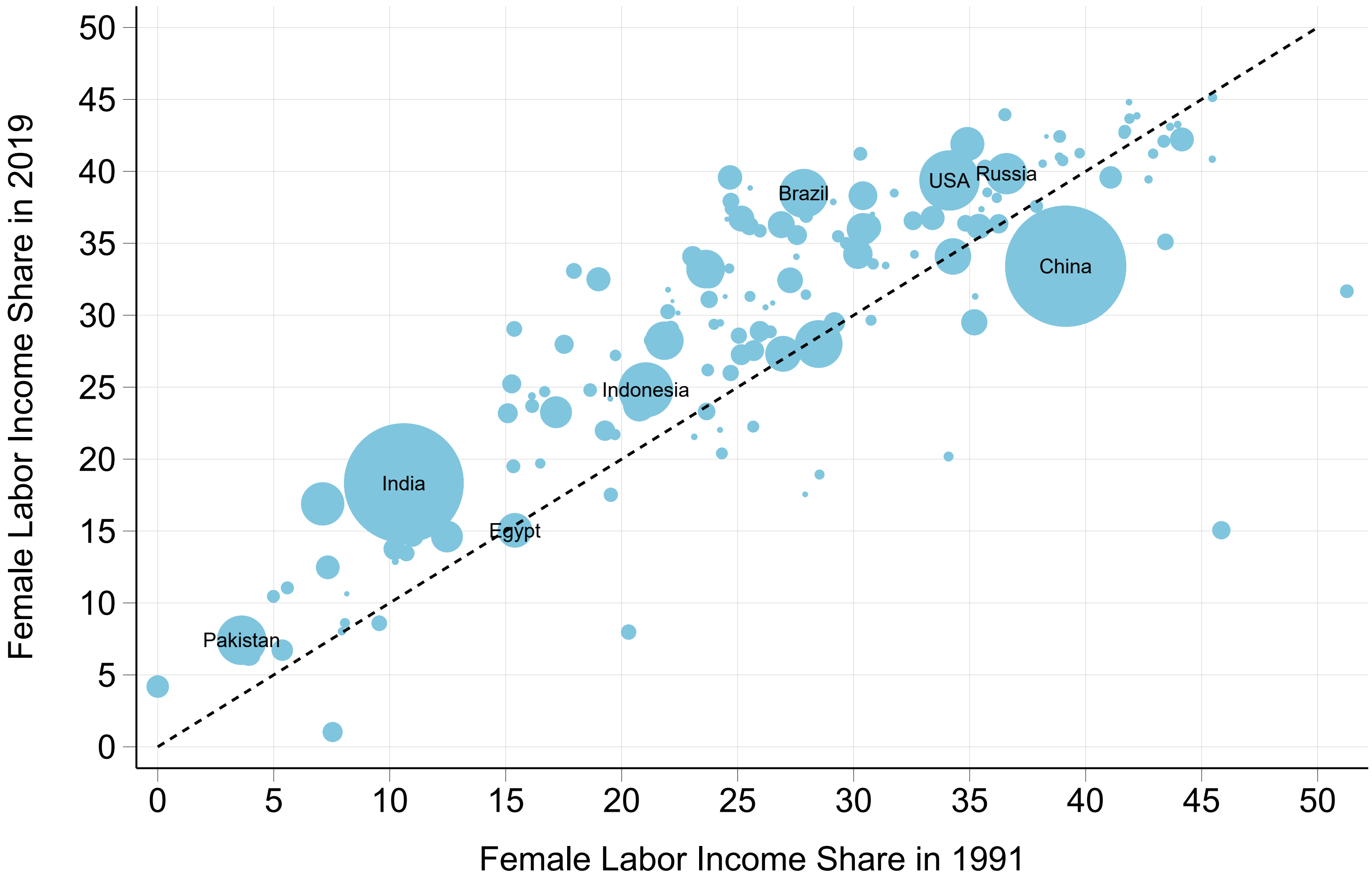
Table A12 – Public Policies and Global Poverty Reduction:
Combining Direct Redistribution and Indirect Investment Benefits from Education (World Bank Data)

	1980	2019	Change (%)	Total Share of Change Explained (%)
Global Poverty Rate ($2.15/Day)				
Pretax Income Absent Educational Expansion	46%	24%	-48%	
Pretax Income	46%	9.7%	-79%	
Posttax Income	44%	4.9%	-89%	46%
Global Bottom 20% Average Income ($/Day)				
Pretax Income Absent Educational Expansion	0.7	1.3	+76%	
Pretax Income	0.7	2.1	+194%	
Posttax Income	0.8	2.8	+241%	68%
Global Bottom 50% Average Income ($/Day)				
Pretax Income Absent Educational Expansion	1.3	2.2	+79%	
Pretax Income	1.3	3.8	+203%	
Posttax Income	1.4	5.0	+263%	70%

Notes. The table compares the evolution of global poverty and the average income of the global bottom 20% and bottom 50% under three scenarios. The first one considers the evolution of each indicator if there had been no educational progress since 1980 ("pretax income absent educational expansion"). The second one corresponds to the actual evolution of each indicator in terms of pretax income ("pretax income"). The third one corresponds to the actual evolution of each indicator in terms of posttax income, that is, after removing all taxes and adding all cash and in-kind transfers (see Gethin, 2023). The last column displays the corresponding share of global poverty reduction or real income gains that can be attributed to public policies, combining direct redistribution (moving from pretax to posttax income) and indirect investment benefits from education (moving from "pretax income absent educational expansion" to pretax income), calculated as one minus the ratio of the first row to the third row of the fourth column within each panel. Global poverty rate and real incomes expressed in 2017 PPP USD. Estimates of the world distribution of income constructed from World Bank per-capita income and consumption data. See table 8 for comparable results using data from the World Inequality Database.

A.2. Additional Results: Global Gender Inequality

Figure A9 – Female Labor Income Share, 1991 vs. 2019

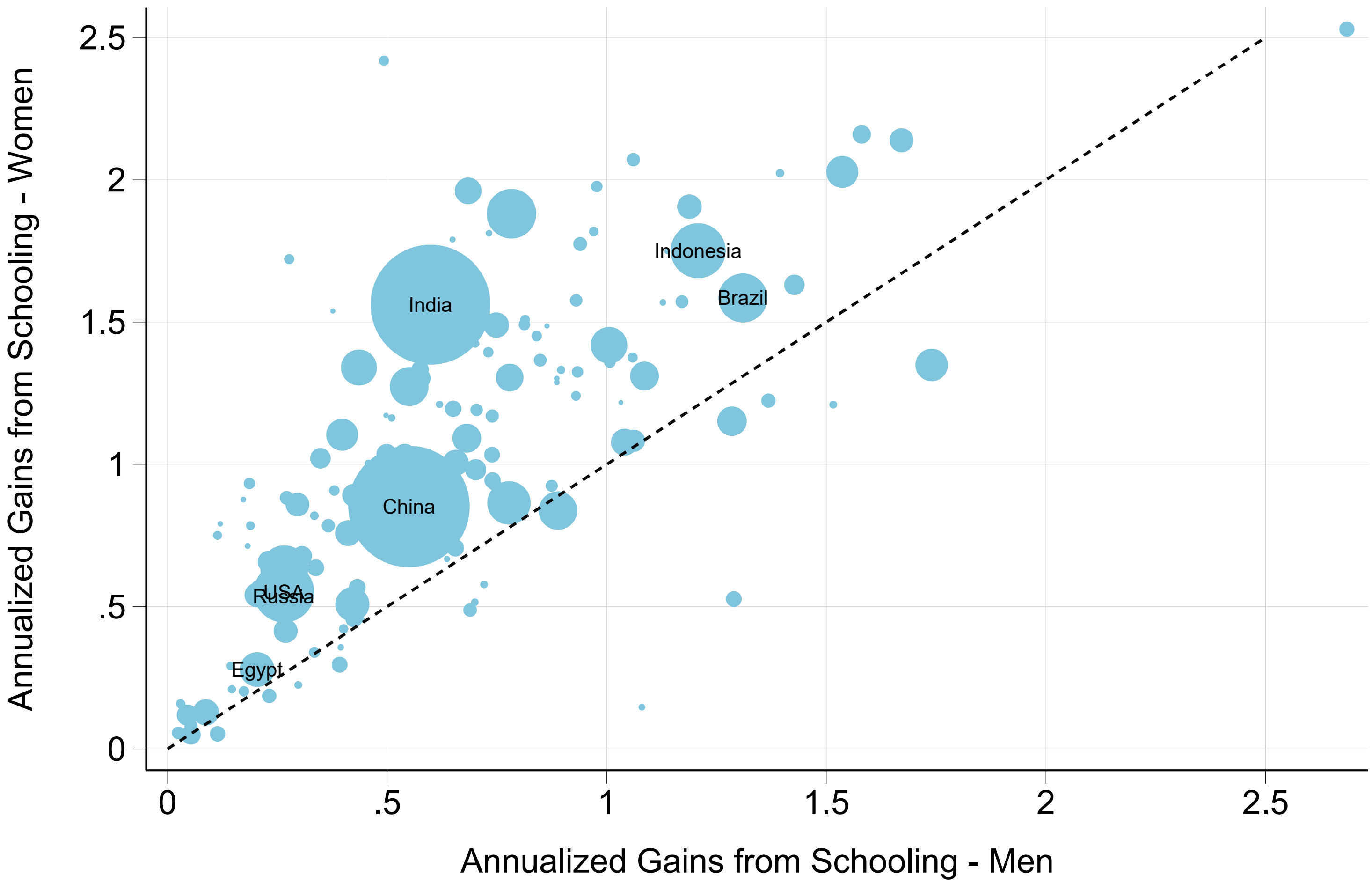


Notes. Data from Neef and Robilliard (2021).

Figure A10 – Returns to Schooling: Men vs. Women

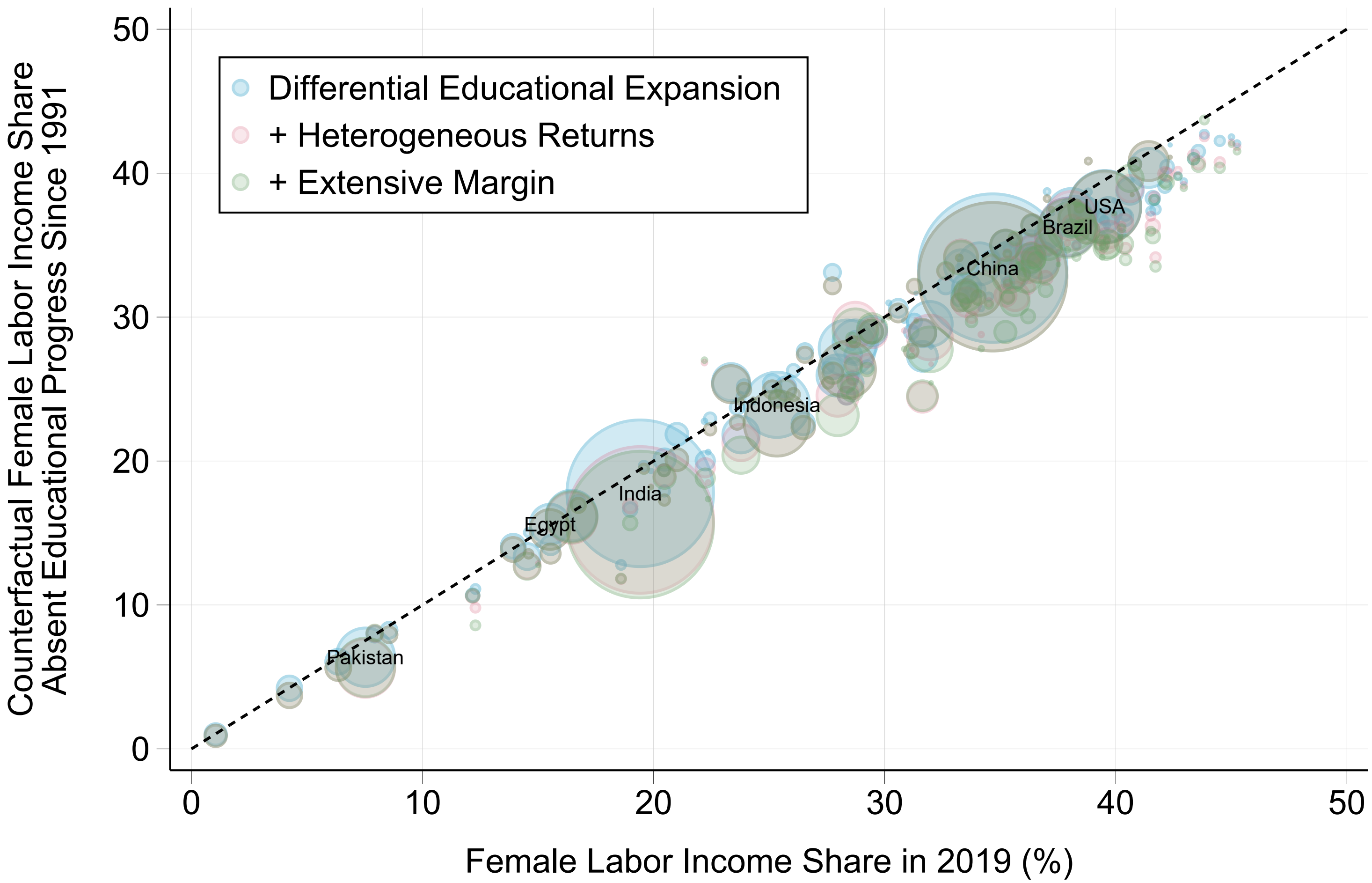
Notes. Author's computations using labor force survey microdata.

Figure A11 – Annualized Income Gains from Schooling, 1991-2019: Men vs. Women



Notes. Author's computations using labor force survey microdata.

Figure A12 – Education and Gender Inequality: Actual vs. Counterfactual Female Labor Income Share in 2019



Notes. Author's computations using labor force survey microdata.

Table A13 – Effect of an Additional Year of Schooling on Female Labor Force Participation: Selected Causal Estimates

Source	Country	Level	β	SE	Baseline	Δ (%)
Akresh, Halim, and Kleemans (2023)	Indonesia	Primary	0		21%	+0%
Khan (2021)	Pakistan	Primary	0		25%	+0%
Grépin and Prashant (2015)	Zimbabwe	Secondary	3	1.7	11%	+27%
Delesalle (2021)	Tanzania	Primary	3.9	1.3		
Cui, Liu, and Zhao (2019)	China	Secondary	4.8	2.7	21%	+23%
Erten and Keskin (2018)	Turkey	Secondary	5		14%	+36%
Spohr (2003)	Taiwan	Secondary	5.8	3.2	45%	+13%
Keats (2018)	Uganda	Primary	7.3	3.1	79%	+9%
Elsayed and Shirshikova (2023)	Egypt	Tertiary	8		31%	+26%
Oliobi (2022)	Nigeria	Tertiary	8.7	2.3	65%	+13%
Chicoine (2021)	Ethiopia	Primary	9.3	5.8	33%	+28%
Hicks and Duanc (2023)	Jordan	Secondary	9.6	3.1	31%	+31%
Kim (2023)	Korea	Tertiary	9.8	2.7	41%	+24%
Overall Average			5.8		35%	+19%

Notes. The table reports estimates of the impact of increasing women's education by one year on female labor force participation (FLFP), based on studies relying on various natural experiments generating quasi-random variation in access to schooling. β: estimated effect of an additional year of schooling on labor force participation. SE: standard error. Baseline: overall labor force participation of women for the sample and definition of employment considered. Δ (%): corresponding percent increase in labor force participation per year of schooling.

Table A14 – Education and Global Gender Inequality, 1991-2019: Excluding China

	1991	2019	Diff.	Share Explained By Education	Share Explained (Cross-Country Average)
Global Female Labor Income Share	28.7%	31.7%	3.1		
Counterfactual: No Educational Progress	28.7%	30.2%	1.6	49%	34%
Counterfactual: + Heterogeneous Returns	28.7%	29.7%	1.0	68%	47%
Counterfactual: + Extensive Margin	28.7%	29.4%	0.7	76%	49%

Notes. The table reports actual versus counterfactual global female labor income shares under different assumptions. China is excluded from the analysis. Global female labor income: total share of labor income received by women in the world as a whole. Change in education: only account for differential trends in schooling by gender, applying the same returns to schooling for men and women to build the counterfactual. Heterogeneous returns: account for differential returns by gender. Extensive margin: account for differential effects of schooling on employment by gender. Cross-country average: population-weighted average of the share of gender inequality reduction explained by education in each country.

Table A15 – Education and Global Gender Inequality, 1991-2019: Average Country, Excluding Countries With Rising Gender Inequality

	1991	2019	Diff.	Share Explained By Education
Actual Female Labor Income Share	20.9%	27.3%	6.4	
Counterfactual: No Educational Progress	20.9%	25.9%	5.1	22%
Counterfactual: + Heterogeneous Returns	20.9%	25.1%	4.2	35%
Counterfactual: + Extensive Margin	20.9%	24.8%	3.9	39%

Notes. The table reports actual versus counterfactual female labor income shares under different assumptions. Figures correspond to the population-weighted average of all countries in the world, excluding all countries where the female labor income share declined, but keeping countries in which educational expansion increased gender inequality. Change in education: only account for differential trends in schooling by gender, applying the same returns to schooling for men and women to build the counterfactual. Heterogeneous returns: account for differential returns by gender. Extensive margin: account for differential effects of schooling on employment by gender.

Table A16 – Education and Global Gender Inequality, 1991-2019:
By Specification of Employment Effects

	Share Explained (World)	Share Explained (Cross-Country Average)
No Employment Effect	73%	57%
Benchmark: OLS Employment Effects	80%	59%
Alternative: +4pp. per Year	100%	67%
Alternative: +6pp. per Year	113%	72%
Alternative: +8pp. per Year	127%	76%
Alternative: +15% per Year	119%	70%
Alternative: +20% per Year	132%	72%
Alternative: +25% per Year	145%	74%

Notes. The table reports the share of the decline in gender inequality that can be explained by education, focusing on the global female labor income share (second column) and the average country (third column; population-weighted), depending on assumptions made on the impact of education on female labor force participation. OLS employment effects: effects of schooling on employment estimated by OLS in each country. Alternative: uniform effect of schooling on female labor force participation, either in terms of percentage points or in terms of relative increases in employment, corresponding to the range of quasi-experimental estimates reported in table A13.

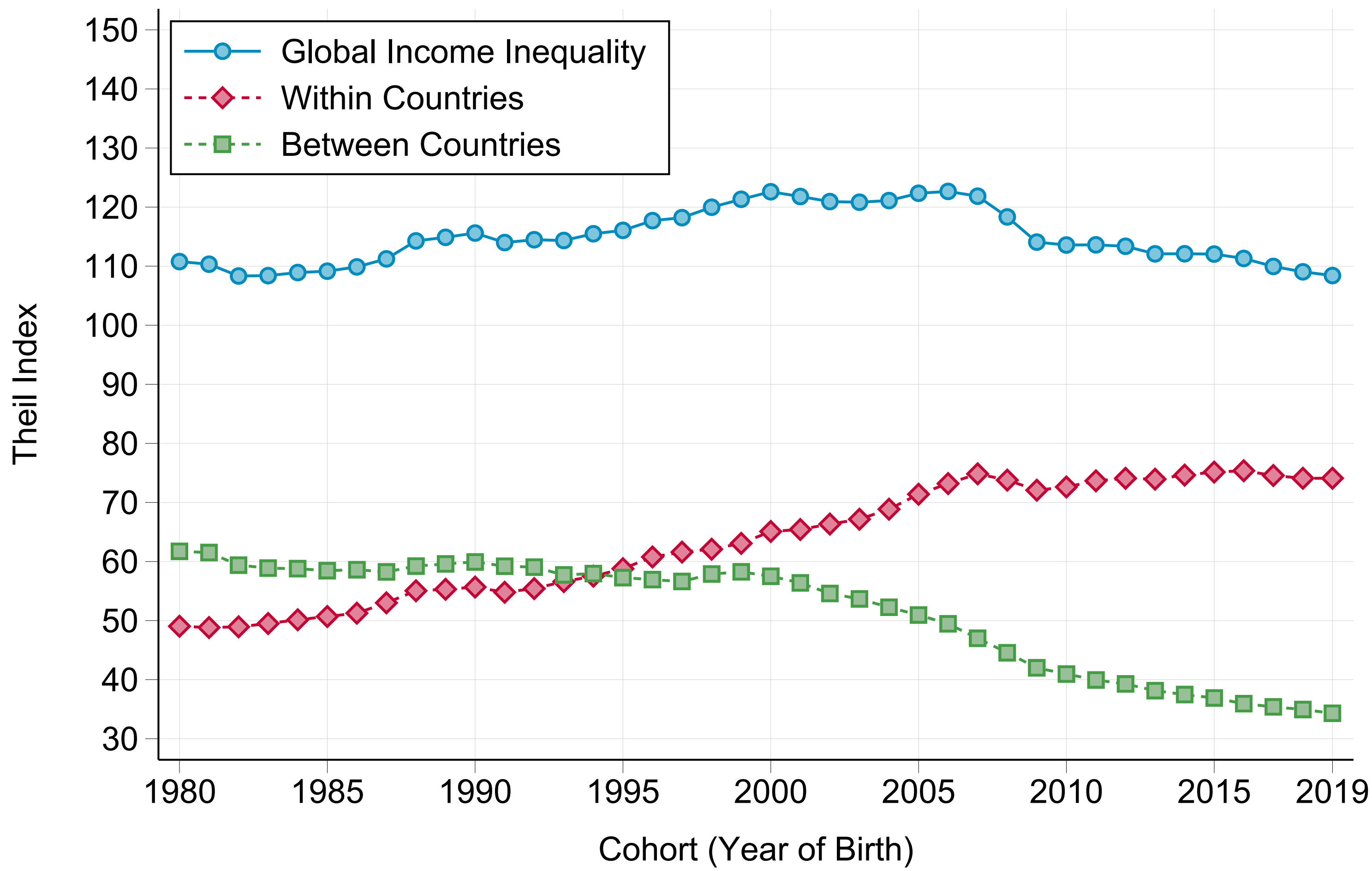
Table A17 – Education and Global Gender Inequality: By Region and Time Period

	Gains from Schooling Ratio Women / Men			Share of Gender Inequality Reduction Explained by Education		
	1991-2019	2000-2019	2010-2019	1991-2019	2000-2019	2010-2019
China	1.5	2.2	1.9	>100%	>100%	>100%
Europe / U.S.	1.9	1.6	2.5	53%	35%	54%
India	2.6	4.7	4.3	47%	49%	68%
Latin America	1.8	1.9	2.1	43%	44%	45%
MENA	2.2	2.3	2.3	64%	73%	49%
Other Asia-Pacific	1.7	1.7	1.8	44%	57%	47%
Sub-Saharan Africa	1.6	1.6	1.6	54%	47%	42%
World Average	1.9	2.4	2.4	59%	58%	61%

Notes. The table reports relative gains from schooling by gender together with the share of gender inequality reduction explained by education for various world regions and time periods. Gains from schooling ratio: women-to-men ratio of annualized income gains from schooling. All numbers correspond to population-weighted cross-country averages of the corresponding indicators in each region. Estimates account for differential educational expansion, heterogeneous returns to schooling, and extensive margin effects.

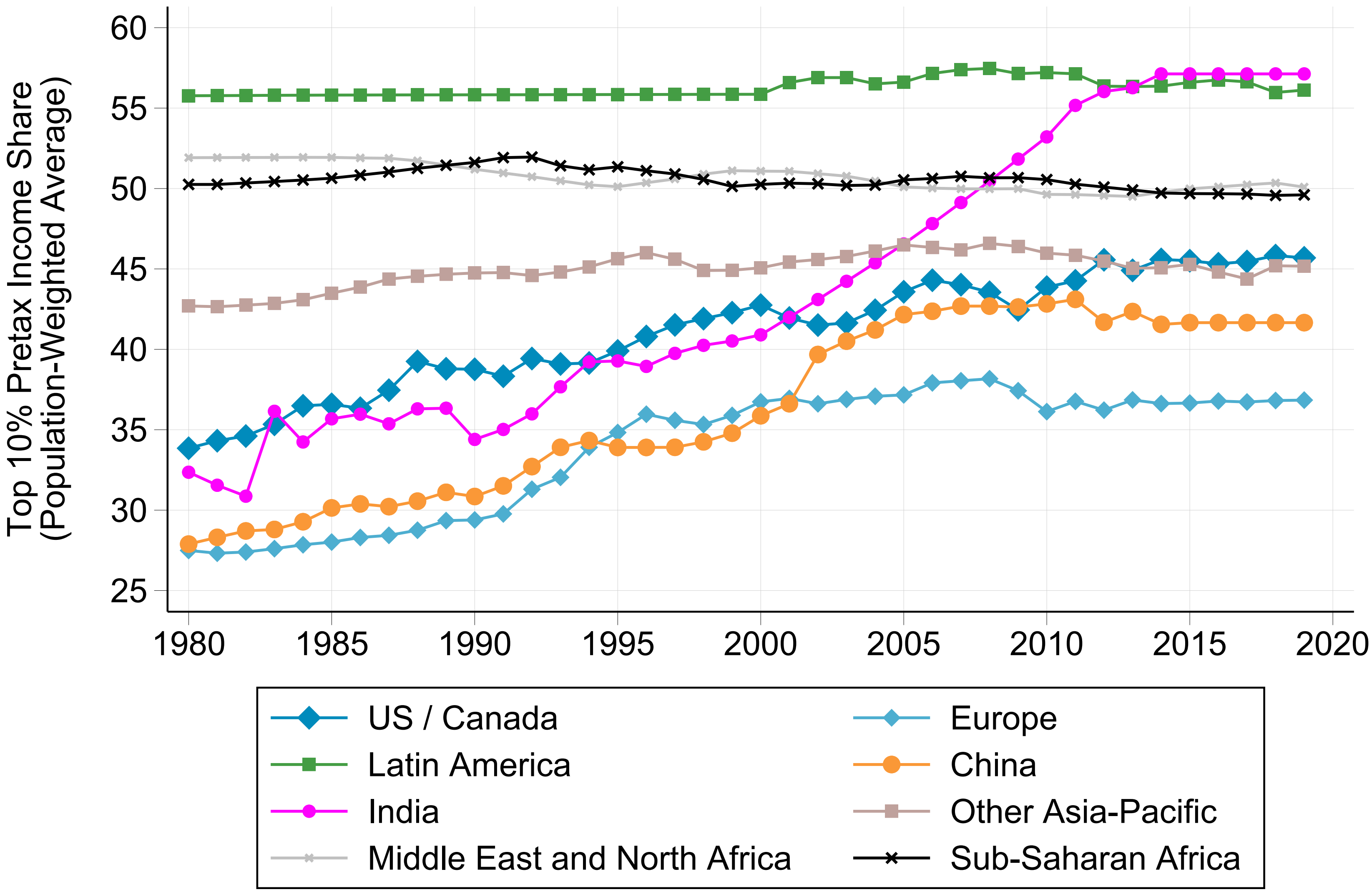
A.3. Stylized Facts on Educational Attainment and the World Distribution of Income

Figure A13 – The Growing Importance of Income Inequality Within Countries
Theil Decomposition of Global Income Inequality, 1980-2019



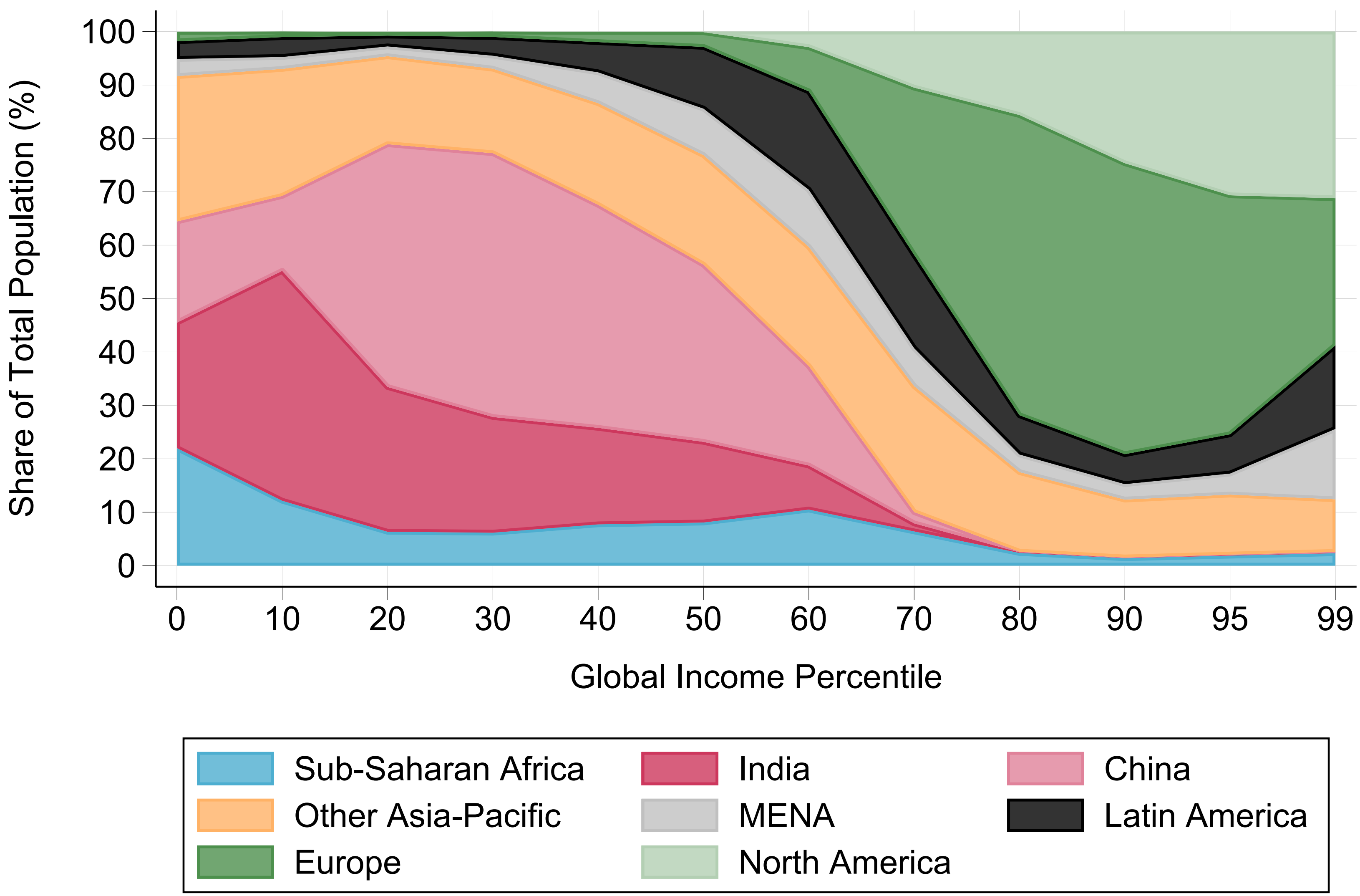
Source: Author's computations using data from the World Inequality Database. The figure plots the evolution of the Theil index of global income inequality from 1980 to 2019, as well as its decomposition into a between-country component and a within-country component.

Figure A14 – The Rise of Within-Country Inequality: Average Top 10% Pretax Income Share by World Region, 1980-2019



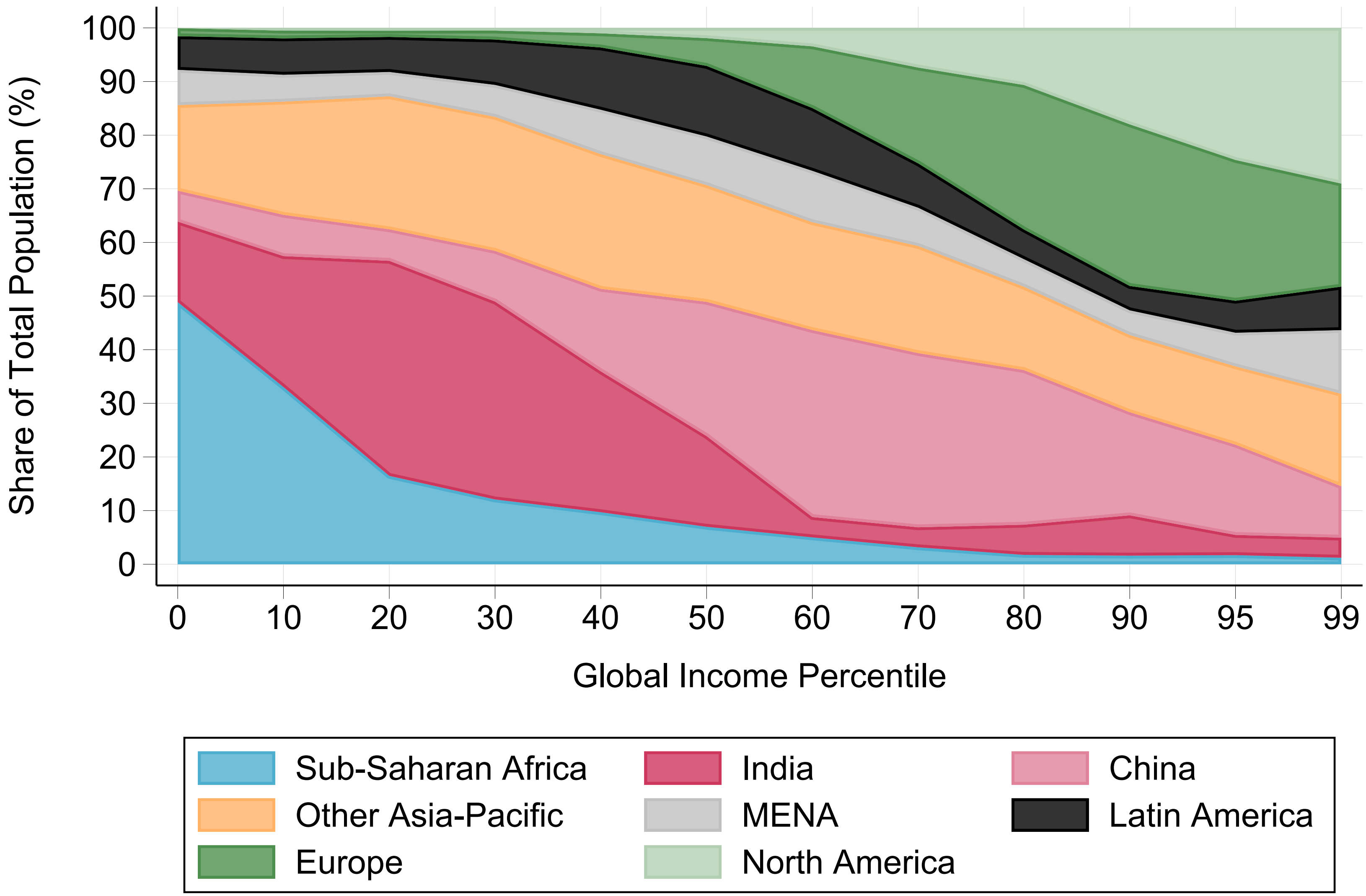
Notes. Author's computations using data from the World Inequality Database.

Figure A15 – Geographical Breakdown of Global Income Groups, 1980



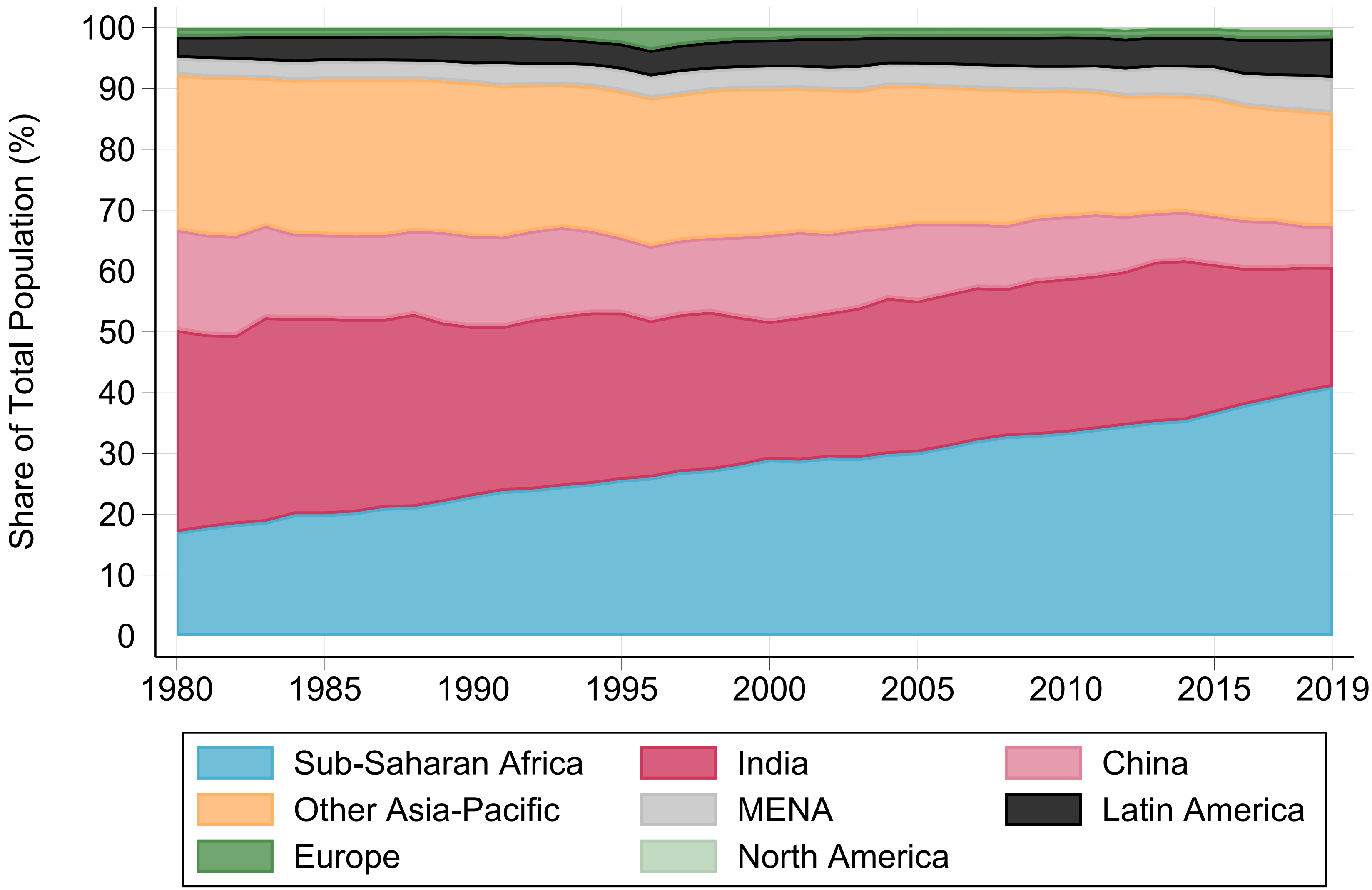
Notes. Author's computations using data from the World Inequality Database.

Figure A16 – Geographical Breakdown of Global Income Groups, 2019



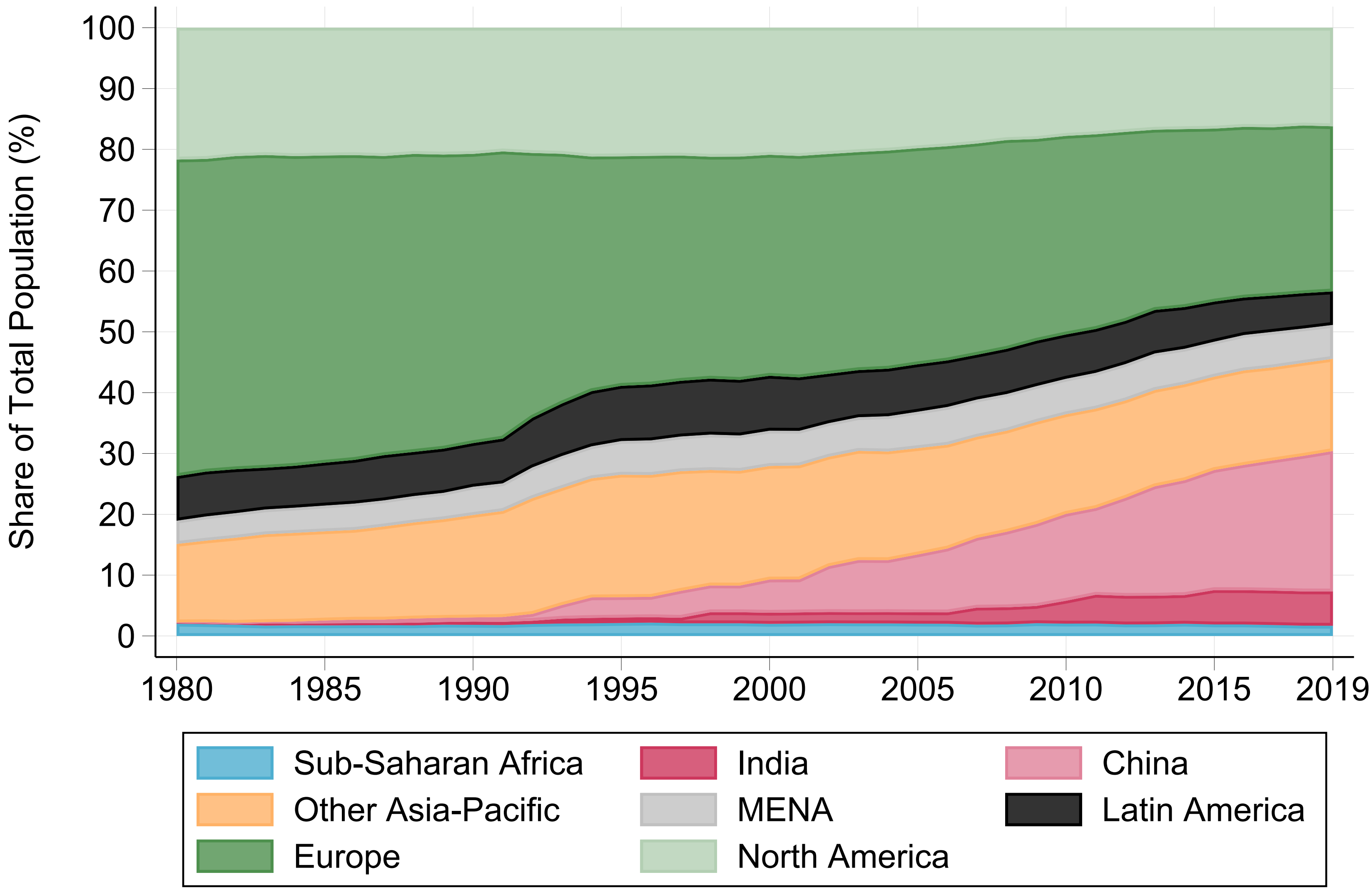
Notes. Author's computations using data from the World Inequality Database.

Figure A17 – Geographical Breakdown of the Global Bottom 20%, 1980-2019



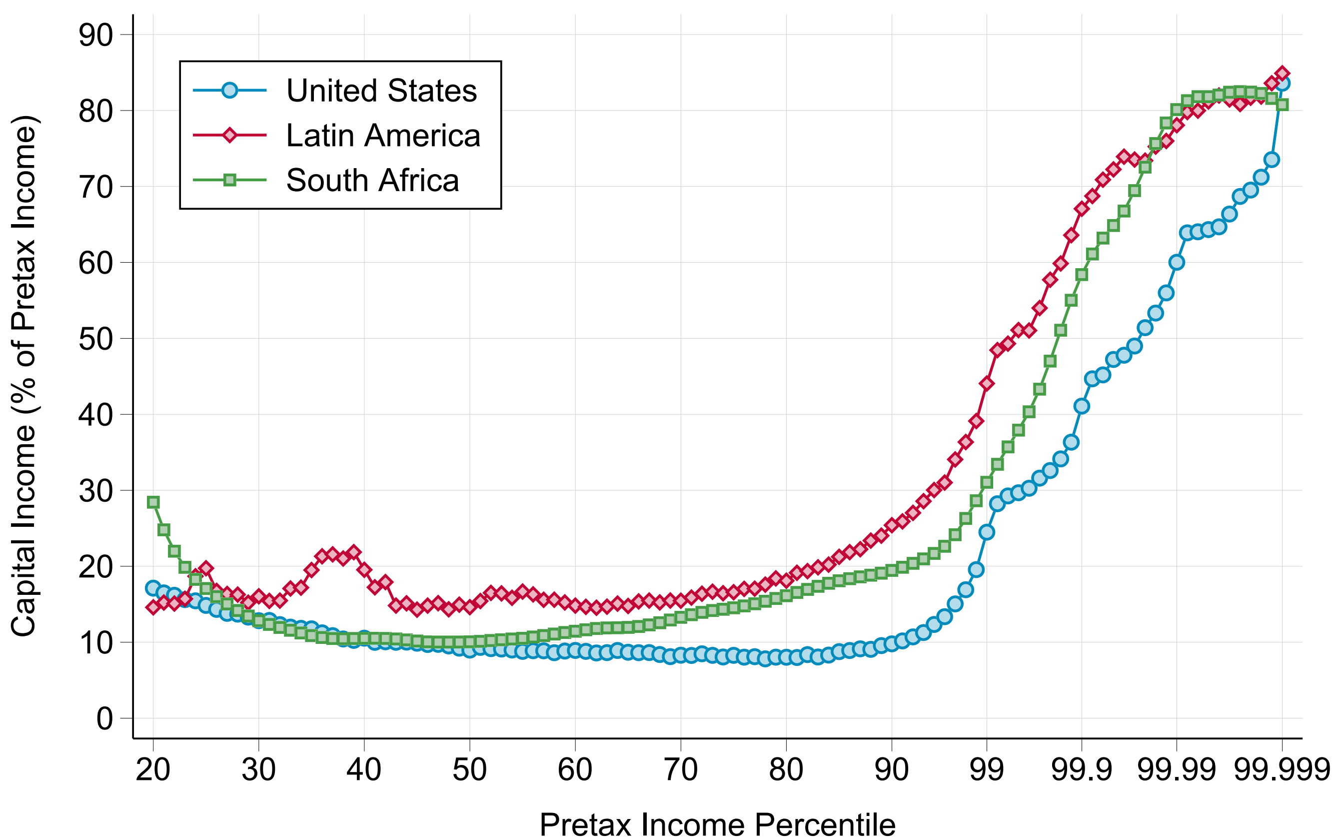
Notes. Author's computations using data from the World Inequality Database.

Figure A18 – Geographical Breakdown of the Global Top 20%, 1980-2019



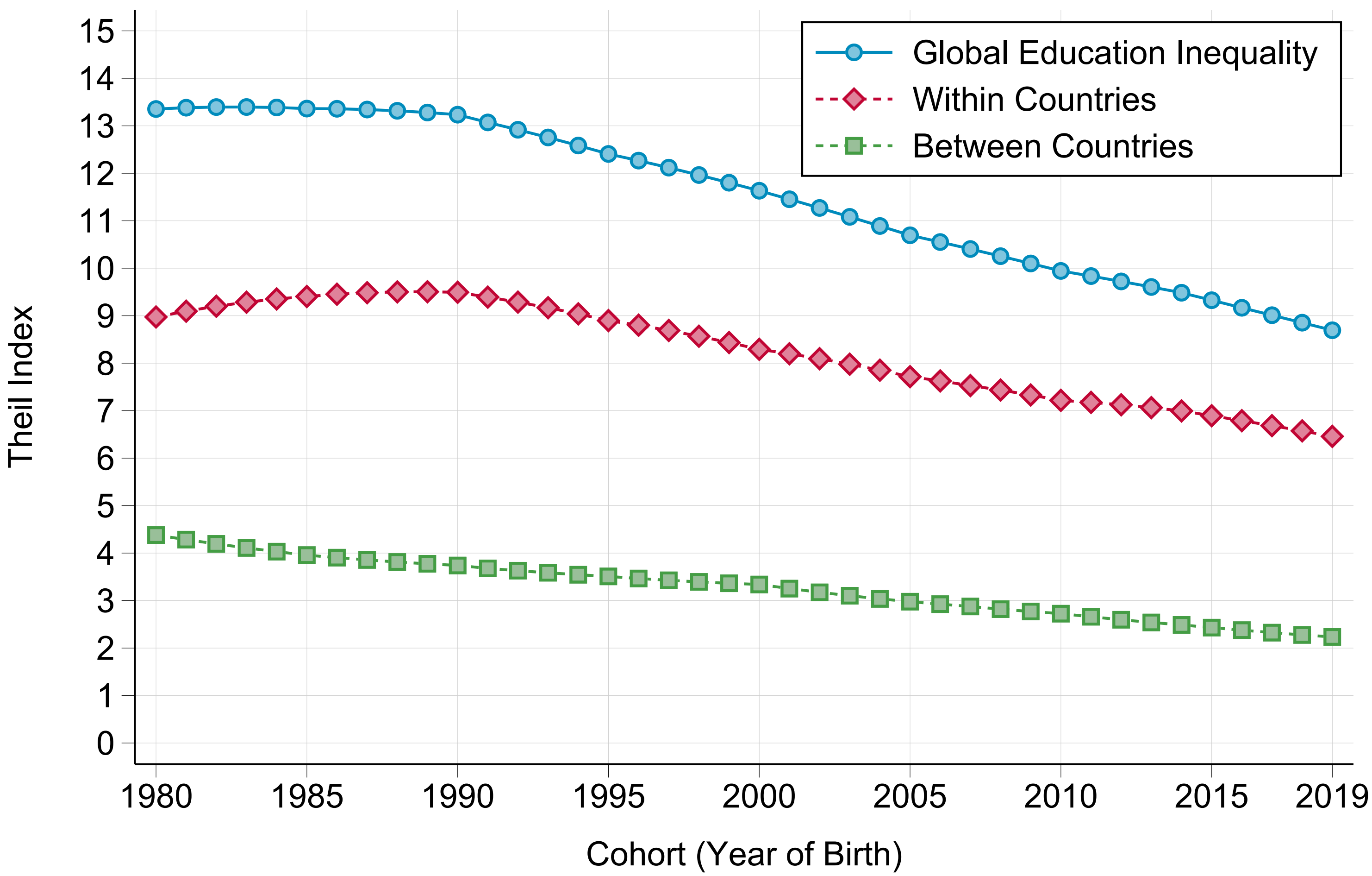
Notes. Author's computations using data from the World Inequality Database.

Figure A19 – The Concentration of Capital Income
in the United States, Latin America, and South Africa



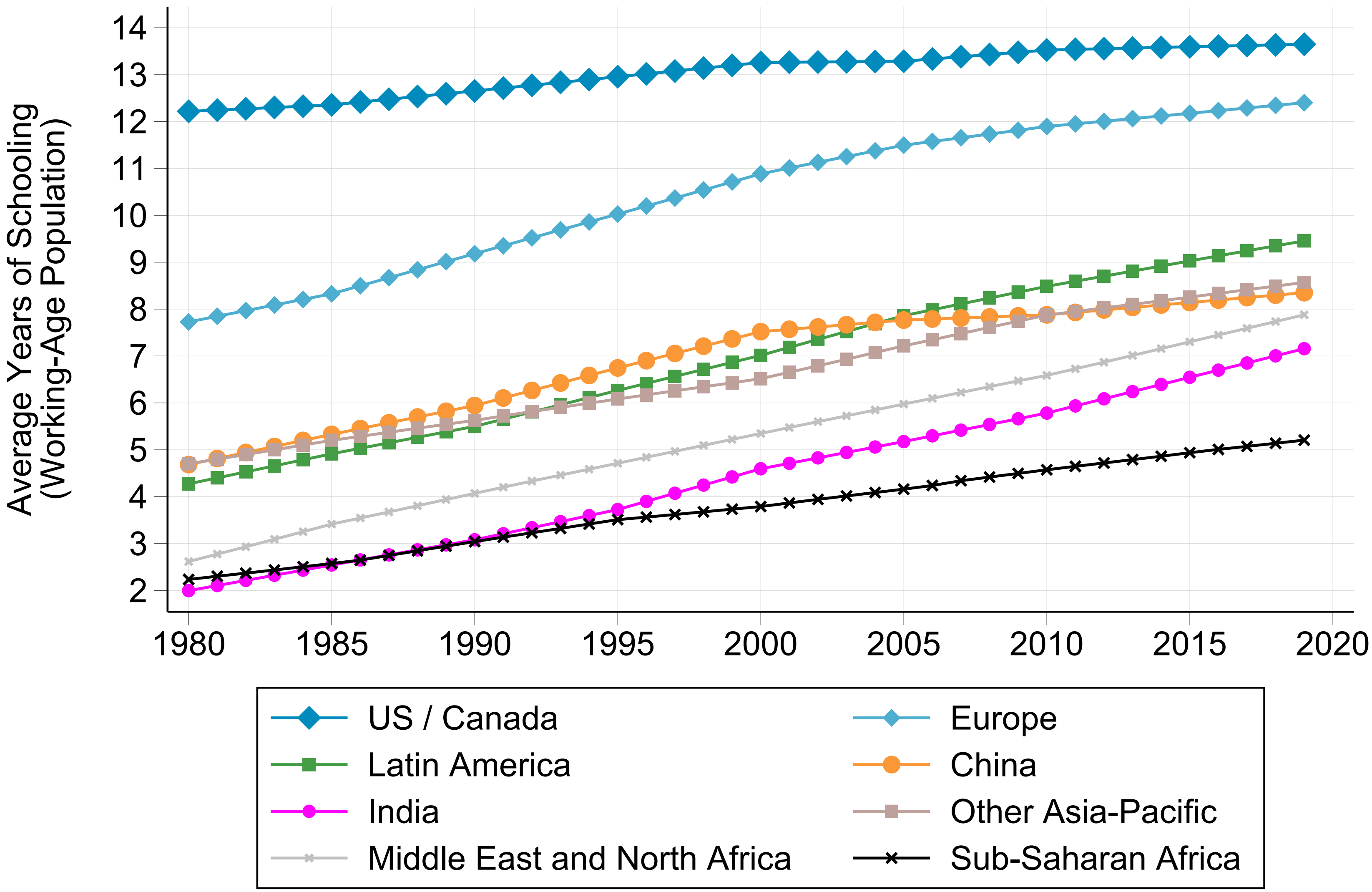
Source: author's elaboration combining data from Piketty, Saez, and Zucman (2018) for the United States, De Rosa, Flores, and Morgan (2022) for Latin America, and Chatterjee, Czajka, and Gethin (2021) for South Africa.

Figure A20 – The Decline of Global Educational Attainment Inequality
Theil Decomposition of Global Human Capital Inequality, 1980-2019



Source: Author's computations combining data from Barro and Lee (2013), updates, and other sources. The figure plots the evolution of the Theil index of global human capital inequality, as well as its decomposition into a between-country component and a within-country component. Human capital at time t in country c with average years of schooling s_{ct} is computed as $H_{ct} = e^{rs_{ct}}$, with r the returns to schooling (set to 10%): see Morrisson and Murtin (2013). This indicator is analogous to the standard deviation of years of schooling.

Figure A21 – Average Years of Schooling by World Region



Notes. Author's computations combining data from Barro and Lee (2013), updates, and other sources.

B. Estimation Details

B.1. Estimation Steps

Sample Restriction The starting point is individual-level data on wages and education. In each survey, I keep individuals aged 25 to 65, with no missing information on age, gender, or education, and with positive reported income.

1) Downgrade Education Levels Next, I match the microdata with information on the distribution of educational attainment by age-gender cell in 1980, covering four education levels: no schooling, incomplete or complete basic education, incomplete or complete secondary education, and incomplete or complete tertiary education. To move from observed educational attainment to counterfactual educational attainment, I randomly sample individuals and downgrade their education levels, until matching 1980 totals by age-gender-level cell.

Individuals belonging to closest education categories are given priority in the simulation. For instance, if 20% of individuals in a given age-gender cell had no schooling in 1980, compared to 10% today, I randomly sample 10% of individuals among the primary education group and downgrade their education level to no schooling. When closest education levels do not contain enough individuals (e.g., only 5% of individuals in this survey have primary education), I instead sample individuals from the category above (secondary education in this example). The outcome is a modified survey, in which the distribution of education levels by age-gender cell corresponds to that observed in 1980. This survey contains "unaffected" observations, corresponding to individuals with unchanged education, as well as "treated" individuals whose education has been downgraded. This approach is very similar to the one recently adopted by Hershbein, Kearney, and Pardue (2020) to estimate the distributional effects of expanding access to college in the United States.

2) Reduce Wages Using Returns to Schooling The second step is to reduce the income of "treated" individuals using estimates of returns to schooling. More precisely, consider an individual i with income y_i, whose education level is downgraded from s_2 to s_1. Denote $r_{s_1,s_2} = \ln(w_1) - \ln(w_2)$ the returns to schooling of moving from s_1 to s_2. Then, the counterfactual income of individual i is:

$$\tilde{y}_i = \exp[\ln(y_i) - r_{s_1,s_2}] \tag{33}$$

I use separate estimates of returns to primary education $r_{non,pri}$, returns to secondary education $r_{pri,sec}$, and returns to tertiary education $r_{sec,ter}$. In the main analysis, these returns are estimated in the microdata using a modified Mincerian equation of the form:

$$\ln y_i = \alpha + \beta^{pri} D_i^{pri} + \beta^{sec} D_i^{sec} + \beta^{ter} D_i^{ter} + X_i \beta + \varepsilon_i \tag{34}$$

Which implies that:

$$r_{non,pri} = \beta^{pri} \tag{35}$$
$$r_{pri,sec} = \beta^{sec} - \beta^{pri} \tag{36}$$
$$r_{sec,ter} = \beta^{ter} - \beta^{sec} \tag{37}$$

For individuals whose education is downgraded by several levels, I use the corresponding cumulative returns. For instance, an individual downgraded from secondary education to no schooling will have a counterfactual income given by $\tilde{y}_i = \exp[\ln(y_i) - r_{sec,pri} - r_{pri,non}]$.

As explained in section 3, I consider lower and upper bounds for returns to schooling. The lower bound corresponds to returns to schooling in 2019, estimated using a modified Mincerian equation. Hence, it corresponds to returns to schooling prevailing under the current distribution of educational attainment:

$$\underline{r}_{s_1,s_2} = r_{s_1,s_2}(L) \tag{38}$$

With $L = (L_1, ..., L_m)$ the distribution of educational attainment in 2019. In contrast, the upper bound corresponds to counterfactual returns to schooling that would prevail if, all other things equal, education levels were to come back to their 1980 levels:

$$\bar{r}_{s_1,s_2} = r_{s_1,s_2}(\tilde{L}) \tag{39}$$

With $\tilde{L} = (\tilde{L}_1, ..., \tilde{L}_m)$ the distribution of educational attainment in 1980. These counterfactual returns to schooling are by construction not observed and have to be estimated. Assuming that the production technology is CES and that we know the elasticity of substitution σ_{s_1,s_2} between s_1 and s_2, the upper bound can be calculated as:

$$\bar{r}_{s_1,s_2} = r_{s_1,s_2}(L) - \frac{1}{\sigma_{s_1,s_2}} \Delta \ln\left(\frac{L_2}{L_1}\right) \tag{40}$$

With $\Delta \ln\left(\frac{L_2}{L_1}\right) = \ln\left(\frac{L_2}{L_1}\right) - \ln\left(\frac{\tilde{L}_2}{\tilde{L}_1}\right)$ the change in the relative supply of skilled workers between

1980 and 2019. An increase in educational attainment from 1980 to 2019 will translate into a decrease in returns to schooling, which implies that the counterfactual return absent educational progress would be higher than the one observed: $\bar{r}_{s_1,s_2} > \underline{r}_{s_1,s_2}$.

In practice, I use the three nests of the production function to calculate three counterfactual returns to primary education, secondary education, and tertiary education:

$$\bar{r}_{non,pri} = r_{non,pri}(L) - \frac{1}{\sigma_3}\Delta \ln\left(\frac{L_{pri}}{L_{non}}\right) \tag{41}$$

$$\bar{r}_{pri,sec} = r_{pri,sec}(L) - \frac{1}{\sigma_2}\Delta \ln\left(\frac{L_{sec}}{L_{pri}}\right) \tag{42}$$

$$\bar{r}_{sec,ter} = r_{sec,ter}(L) - \frac{1}{\sigma_1}\Delta \ln\left(\frac{L_{sec}}{L_{ter}}\right) \tag{43}$$

Finally, using the CES production function, it is possible to recover the relative weight that should be put on counterfactual versus observed returns to obtain the true return to schooling (see section 2.2). Figure B2 plots the empirical distribution of these weights across all countries for each of the three nests. The weight put on initial (counterfactual) returns mostly ranges from 0.5 to 0.7 for all three nests, with typical values in-between 0.55 and 0.65. The true return is thus close to the average of observed and counterfactual return, with a slightly greater weight given to the latter. Figures B3, B4, and B5 display the corresponding observed versus true returns to primary, secondary, and tertiary education in each country.

3) Adjust Relative Wages The third step is to adjust relative wages of both unaffected and treated individuals, using the nested CES specification presented in section 3. This step of the estimation leaves the average income in the survey unchanged, since aggregate effects are captured in the previous step of the estimation. It then suffices to adjust relative wages using the three elasticities of substitution, while keeping average income constant. In practice, I do this in three steps in the microdata.

First, I reduce the log average income of primary-educated workers by the product of the primary/no schooling supply shift and the elasticity of substitution σ_3, and readjust earnings within this nest ($L_{\bar{sec}}$) to leave the average unchanged:

$$\Delta \log\left(\frac{w_{pri}}{w_{non}}\right) = -\frac{1}{\sigma_3}\Delta \log\left(\frac{L_{pri}}{L_{non}}\right) \tag{44}$$

Second, I repeat the same procedure for the secondary/below-secondary nest ($L_{\bar{t}er}$):

$$\Delta \log\left(\frac{w_{sec}}{w_{s\bar{e}c}}\right) = -\frac{1}{\sigma_2}\Delta \log\left(\frac{L_{sec}}{L_{s\bar{e}c}}\right) \tag{45}$$

Third, I repeat the same procedure for the upper level of the CES production function:

$$\Delta \log\left(\frac{w_{ter}}{w_{t\bar{e}r}}\right) = -\frac{1}{\sigma_1}\Delta \log\left(\frac{L_{ter}}{L_{t\bar{e}r}}\right) \tag{46}$$

In a handful of countries, the share of the working-age population with no schooling or primary education declined to almost zero in 2019, leading some relative supply shifts to diverge to infinity. To avoid this divergence, I bound the absolute value of changes in relative supply to 4 log points. This does not affect the results significantly, given that concerned countries are those where the initial level of low-skilled workers was already very small.

4) Growth Accounting The final step is to use this counterfactual to estimate the share of growth explained by education. I first aggregate actual labor income Y_L and counterfactual labor income $\tilde{Y}_L$ in the survey microdata by decile, and calculate the corresponding ratio of counterfactual income to actual income: $\psi^d = \frac{\tilde{Y}_L^d}{Y_L^d}$. This yields a measure of how much lower labor income would be if education had not improved.

I then incorporate these estimates into global income distribution data. I start with distributions from the World Inequality Database, which provide information on the average pretax income of each generalized percentile (all percentiles from p0 to p99, followed by a further decomposition of top incomes up to p99.999p100). I merge estimates of ψ^d by country-year-decile with this database.[27] I then calculate counterfactual total pretax income of generalized percentile p as:

$$\tilde{Y}^p = Y_K^p + \psi^p Y_L^p \tag{47}$$

Finally, I construct separate actual and counterfactual world distributions of income from 1980 to 2019, by ranking all individuals in the world by each income concept and aggregating average income by global generalized percentile.

[27]To get smoother profiles of counterfactual income by generalized percentile, I assume that ψ^d for each decile corresponds to the ratios observed for p5, p15, p25, p35, p45, p55, p65, p75, p85, and p95. I then interpolate ψ^d between percentiles to fill in missing values. I assume that values observed for percentiles within the bottom 5% and the top 5% are those observed for p5 and p95, respectively. Finally, I drop generalized percentiles with zero average pretax income in the World Inequality Database.

B.2. Aggregate and Individual Returns to Schooling: CES Simulations

B.2.1. Theoretical Background

The objective of this section is to shed light on the following question: which returns to schooling should be used to estimate the effect of educational expansion on total output? And how far is this return from the return observed before (initial return) versus after (final return) increasing education? Consider a CES production function with two skill types:

$$Y = \left(A_H L_H^{\frac{\sigma-1}{\sigma}} + A_L L_L^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}$$

We normalize $L_H + L_L = 1$, so that L_H corresponds to the share of skilled workers in the economy. The objective is to compare output under an initial distribution of skills $\{L_{H1}, L_{L1}\}$ and a final distribution of skills $\{L_{H2}, L_{L2}\}$, with $L_{H2} > L_{H1}$.

One possibility is to predict the change in output using initial returns to schooling r_1:

$$r_1 = \log\left(\frac{w_{H1}}{w_{L1}}\right) = \frac{\sigma-1}{\sigma}\log\left(\frac{A_H}{A_L}\right) - \frac{1}{\sigma}\log\left(\frac{L_{H1}}{L_{L1}}\right)$$

Predicted output can then be calculated as a weighted average of the wages of always skilled workers, always unskilled workers, and newly skilled workers:

$$Y = w_{H1}L_{H1} + w_{L1}L_{L2} + w_{H1}(L_{L1} - L_{L2})$$

Where $w_{H1} = \exp\left(\log(w_{L1}) + r_1\right)$ is the wage of high-skilled workers in the initial period. This amounts to considering that supply effects change relative wages, but do not reduce the effect of educational expansion on output. Initial returns to schooling then capture exactly the effect of skill upgrading. As demonstrated by Caselli and Ciccone (2013), this is an upper bound under standard assumptions on the production technology (which are satisfied in the CES case). This is because supply effects decrease the marginal product of skilled workers and increase that of unskilled workers, but the former effect dominates the latter.

An alternative possibility is to use final returns to schooling r_2:

$$r_2 = \log\left(\frac{w_{H2}}{w_{L2}}\right) = \frac{\sigma-1}{\sigma}\log\left(\frac{A_H}{A_L}\right) - \frac{1}{\sigma}\log\left(\frac{L_{H2}}{L_{L2}}\right)$$

Predicted output can then be calculated as:

$$Y = w_{H1}L_{H1} + w_{L1}L_{L2} + \exp\left(\log(w_{L1}) + r_2\right)\left(L_{L1} - L_{L2}\right)$$

This amounts to assuming that supply effects reduce the effect of educational expansion on output by the same amount as the decrease in returns to schooling. The benefits of skill upgrading then correspond exactly to the returns to schooling observed at the end of the period. This constitutes a lower bound on the actual effect of educational expansion, which may underestimate it significantly. In particular, consider the extreme case in which a large shock to the supply of skilled workers would bring returns to schooling to zero. This approach would then predict no change in output from educational expansion, which is impossible in this model as long as we assume that $A_H > A_L$.

B.2.2. Simulation

To investigate which weight should be put on final versus initial returns to schooling, I simulate predicted and actual changes in output under different parametrizations of the production function. More specifically, I run simulations jointly varying the elasticity of substitution σ from 1.5 to 8, the relative efficiency of skilled workers A_H/A_L from 1.2 to 5, and the final share of skilled workers L_{H2} from 0.15 to 0.95 (assuming an initial value $L_{H1} = 0.1$).

Given that all parameters are specified, it is also possible to calculate the actual individual return r^* that should be used to predict changes in output. This return satisfies:

$$\begin{aligned} Y &= w_{H1}L_{H1} + w_{L1}L_{L2} + \exp\left(\log(w_{L1}) + r^*\right)\left(L_{L1} - L_{L2}\right) \\ \Rightarrow r^* &= \log\left(\frac{Y - w_{H_1}L_{H1} - w_{L_1}L_{L1}}{L_{L1} - L_{L2}}\right) - \log(w_{L1}) \end{aligned}$$

Finally, we also observe initial and final returns r_1 and r_2, which means that one can calculate the "optimal weight" γ that should be put on each return:

$$\begin{aligned} r^* &= \gamma r_1 + (1-\gamma) r_2 \\ \Rightarrow \gamma &= \frac{r^* - r_2}{r_1 - r_2} \end{aligned}$$

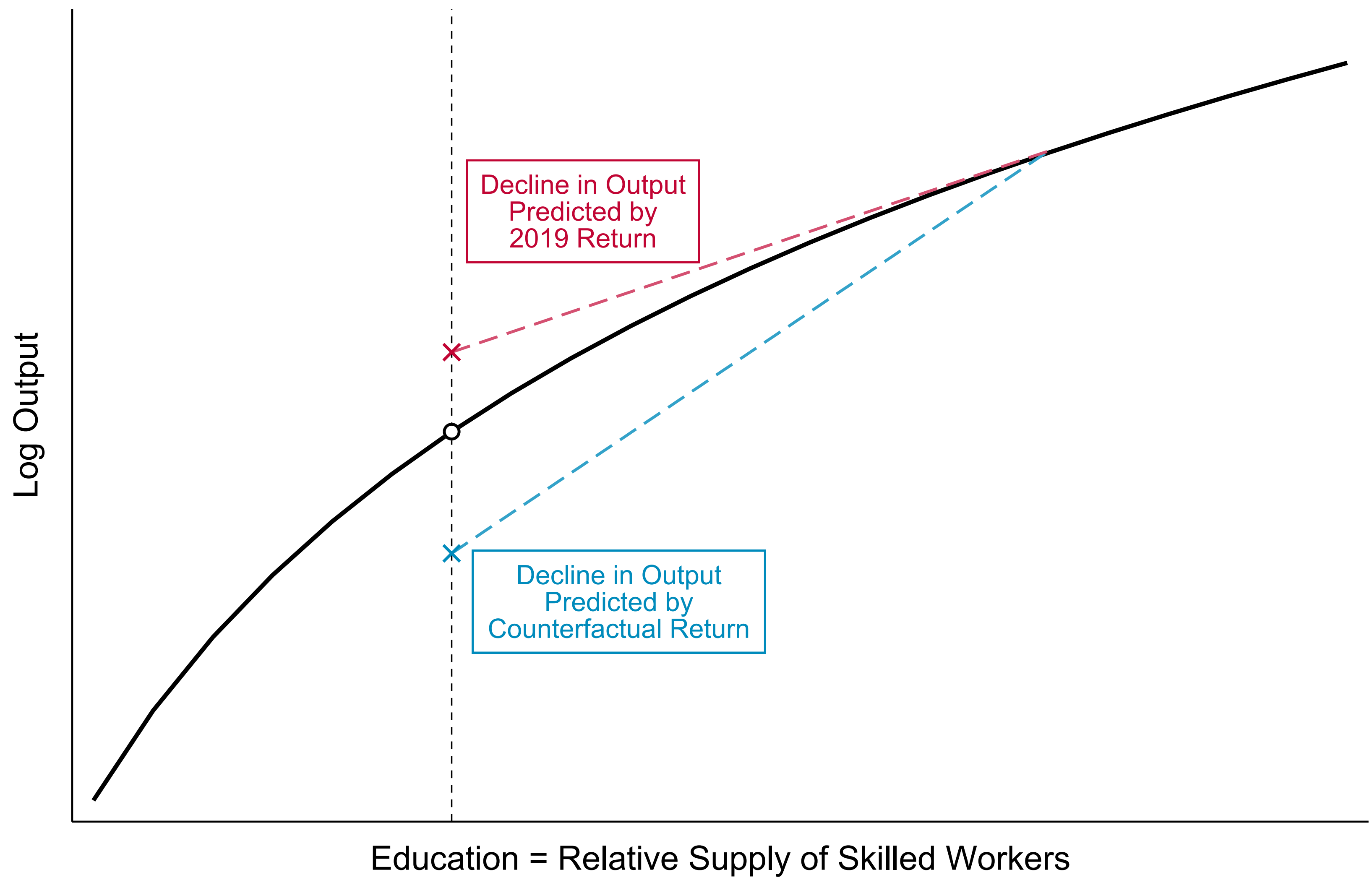
A weight of 0.5 means that the average of initial and final returns provides a good approximation for the true effect of educational expansion on output. A weight greater than 0.5 means that we should give more importance to initial returns in the estimation.

B.2.3. Results

Figure B9 plots the resulting distribution of optimal weights on initial returns across all combinations of parameters. For elasticities ranging from 1.5 to 8 and relative skill efficiencies ranging from 1.2 to 5, the weight on initial returns ranges from about 0.45 to 0.8, with a mean of 0.65. Almost no estimate falls below 0.5, meaning that initial returns are almost always closer to the true effect of educational expansion on output than final returns.

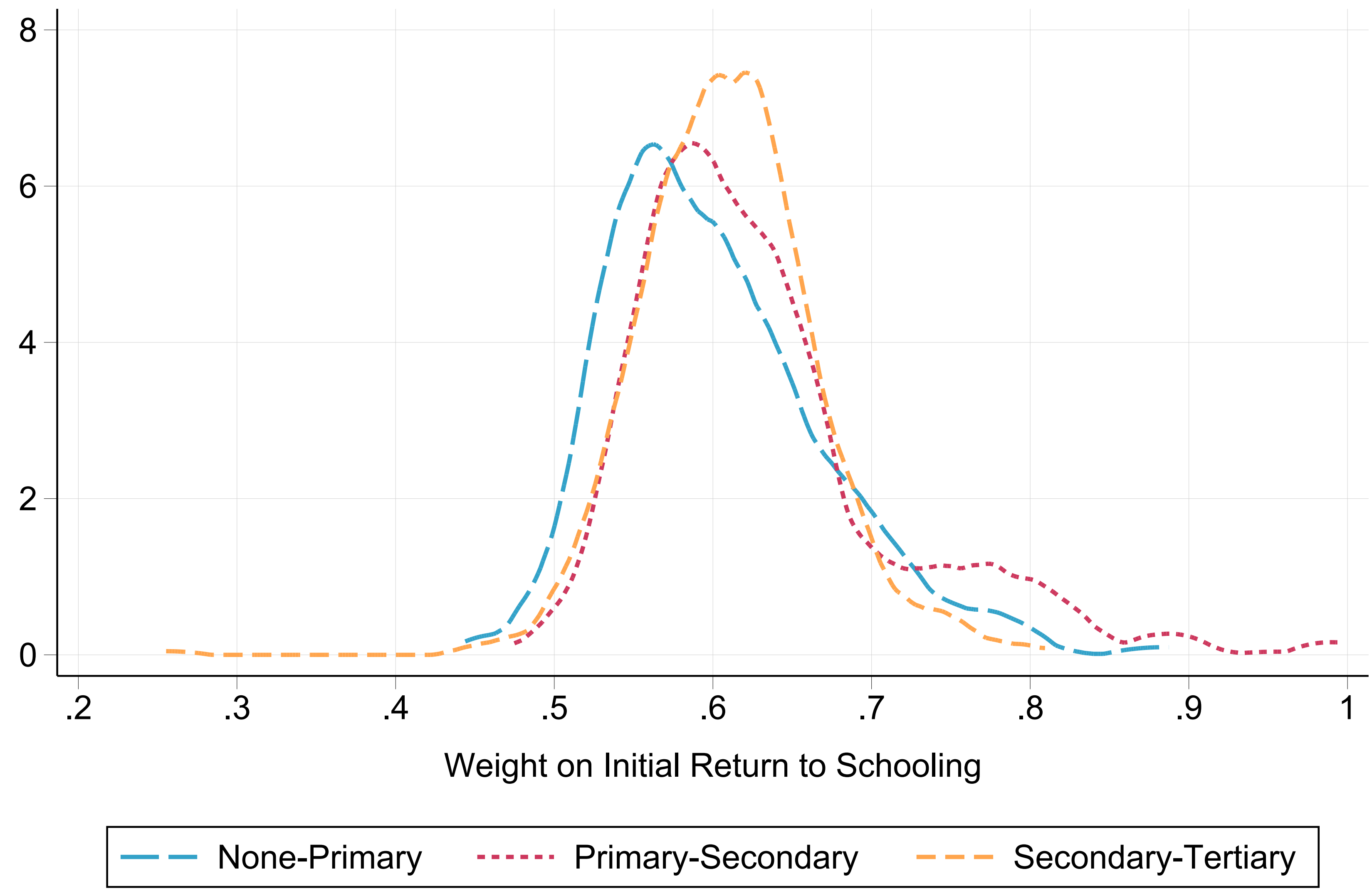
Figure B10 shows how the weight on initial returns varies across specific combinations of relative skill efficiencies and the elasticity of substitution. The weight is higher for lower elasticities of substitution and for higher levels of relative skill efficiencies. For a long-run elasticity of substitution of 4, the weight ranges from about 0.5 in case of very low differences in efficiencies to 0.7 for a scenario in which high-skill workers are 5 times more efficient than low-skill workers.

Figure B1 – Initial, Final, and True Returns to Schooling: Graphical Illustration



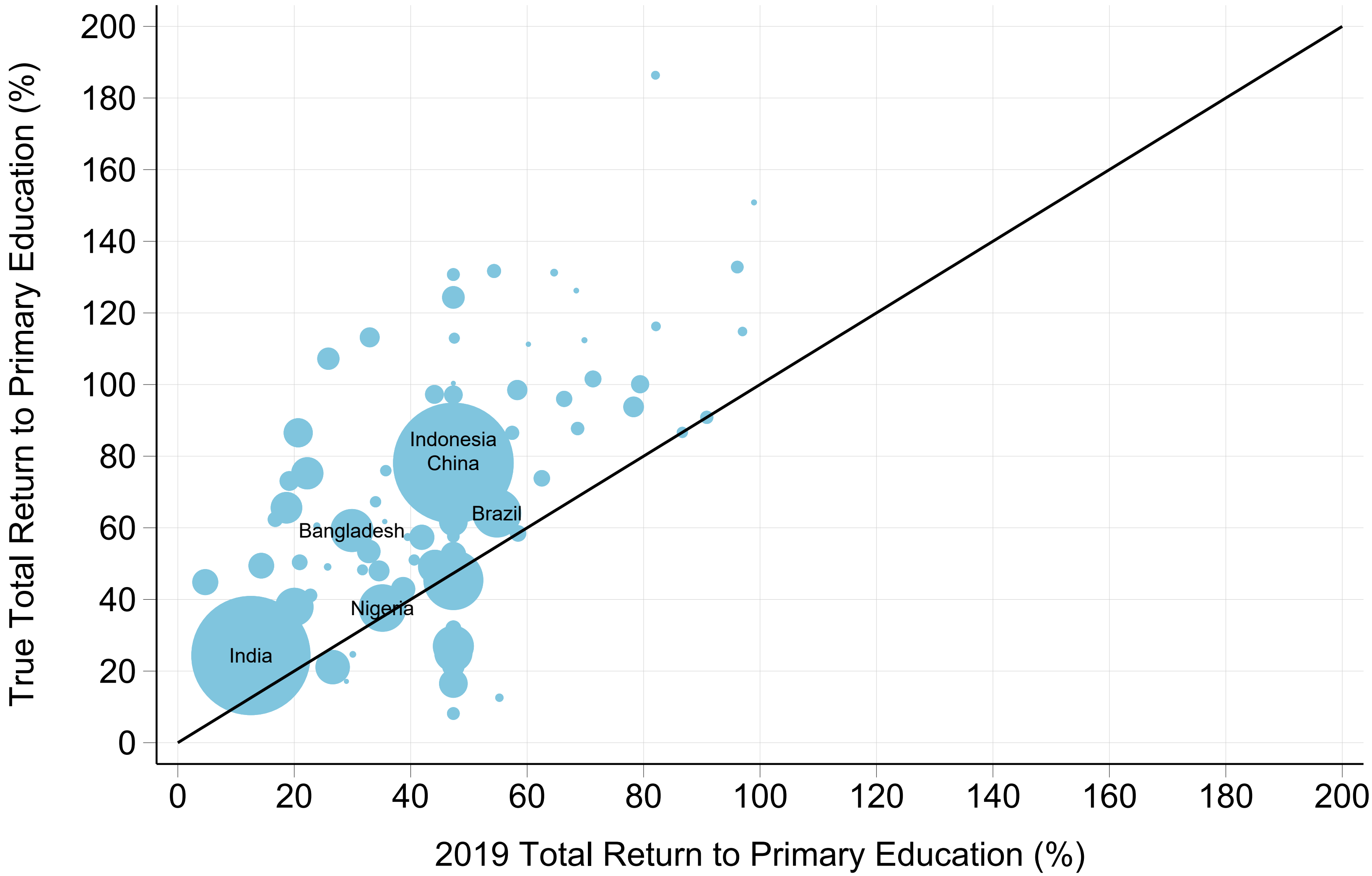
Notes. The figure provides a graphical illustration of why true returns to schooling are in-between initial and final returns. The upper dashed line shows the decline in output from reducing education as predicted by the return to schooling observed in 2019 (corresponding to final returns). The lower dashed line shows the output loss predicted by the return to schooling that prevails after reducing education (corresponding to initial returns). The true decline in output lies in-between the two estimates.

Figure B2 – Returns to Schooling: Empirical Distribution of Optimal Weights on Initial Returns



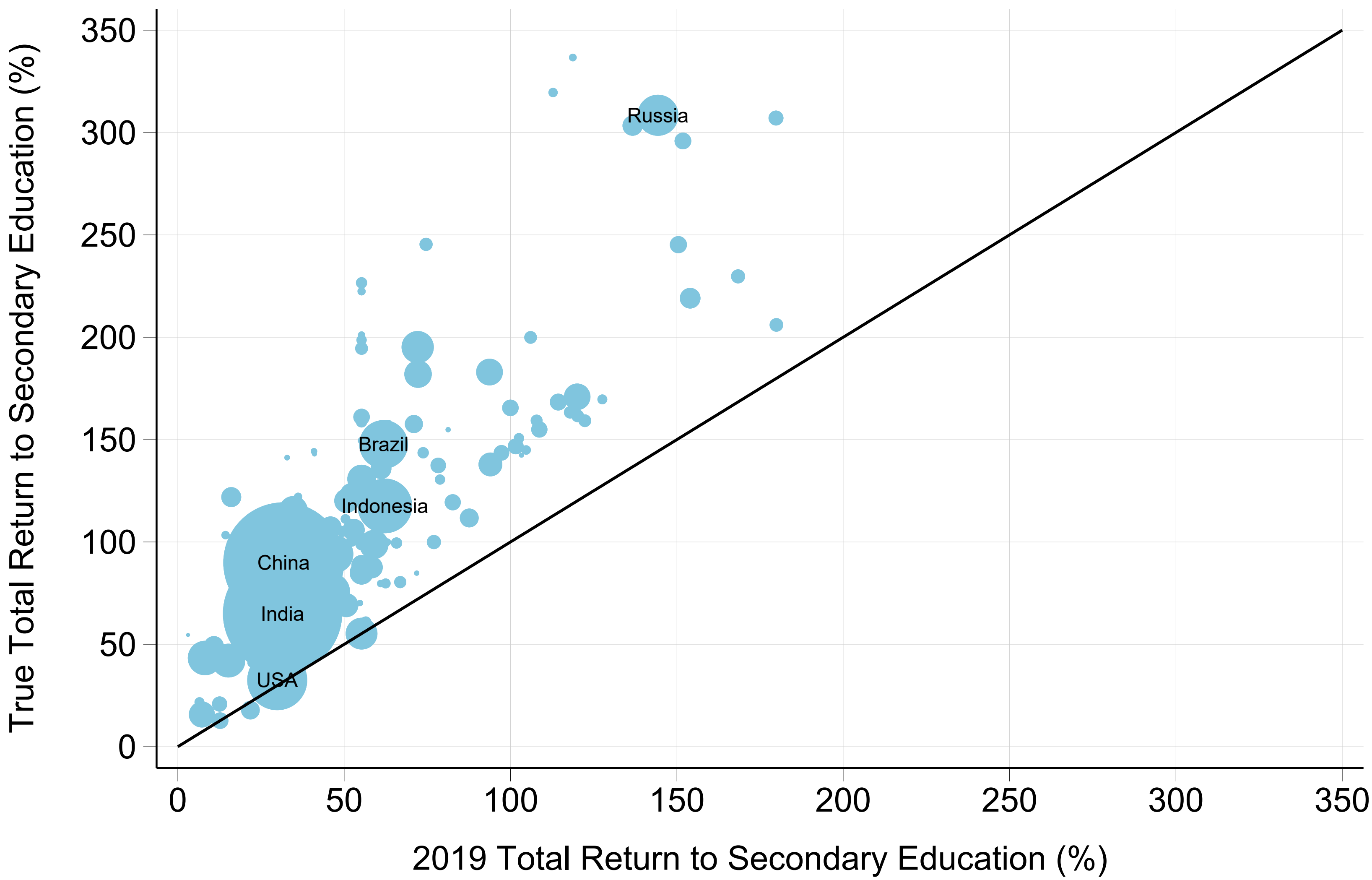
Notes. The figure plots the empirical distribution of optimal weights on initial returns to schooling required to estimate the true effect of educational expansion on output, for each of the three nests of the production function. Estimates assume an elasticity of substitution of 6.

Figure B3 – Returns to Schooling: 2019 vs. True Total Return to Primary Education



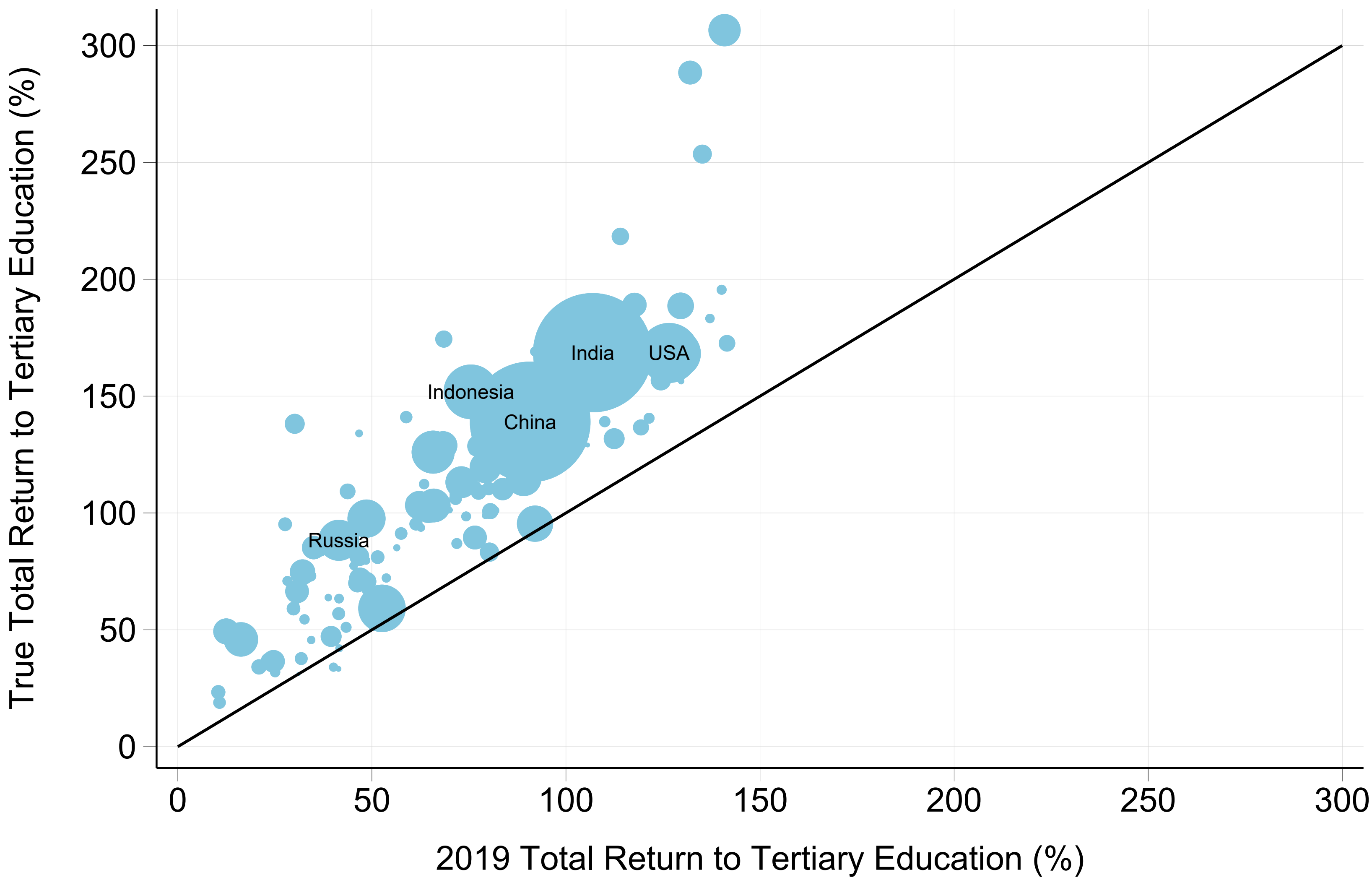
Notes. The figure compares the 2019 (final) and true total return to primary education (the percent increase in personal income of moving from no schooling to primary education) in each country. True returns are estimated assuming an elasticity of substitution of 5.

Figure B4 – Returns to Schooling: 2019 vs. True Total Return to Secondary Education



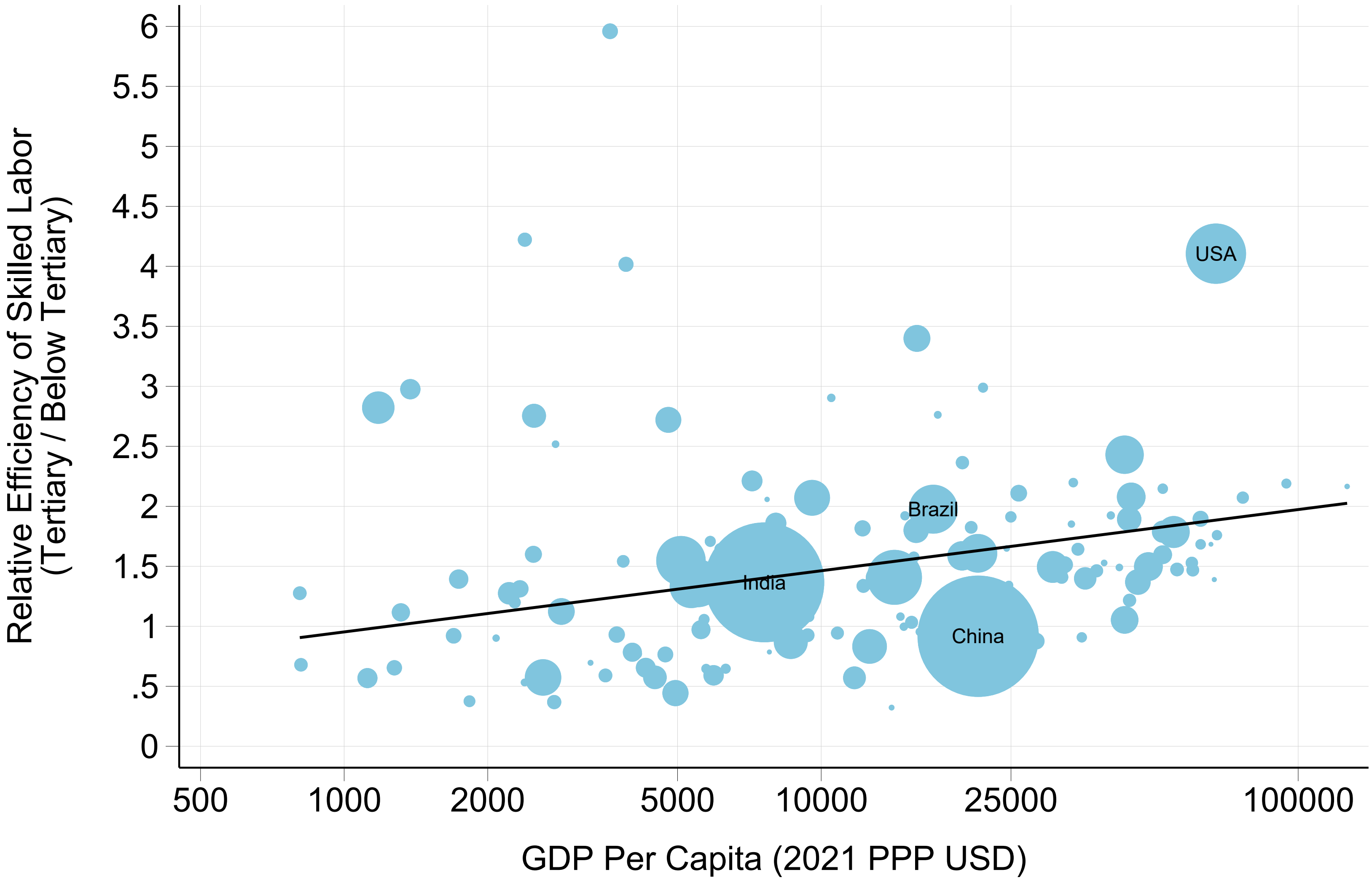
Notes. The figure compares the 2019 (final) and true total return to secondary education (the percent increase in personal income of moving from primary to secondary education) in each country. True returns are estimated assuming an elasticity of substitution of 5.

Figure B5 – Returns to Schooling: 2019 vs. True Total Return to Tertiary Education



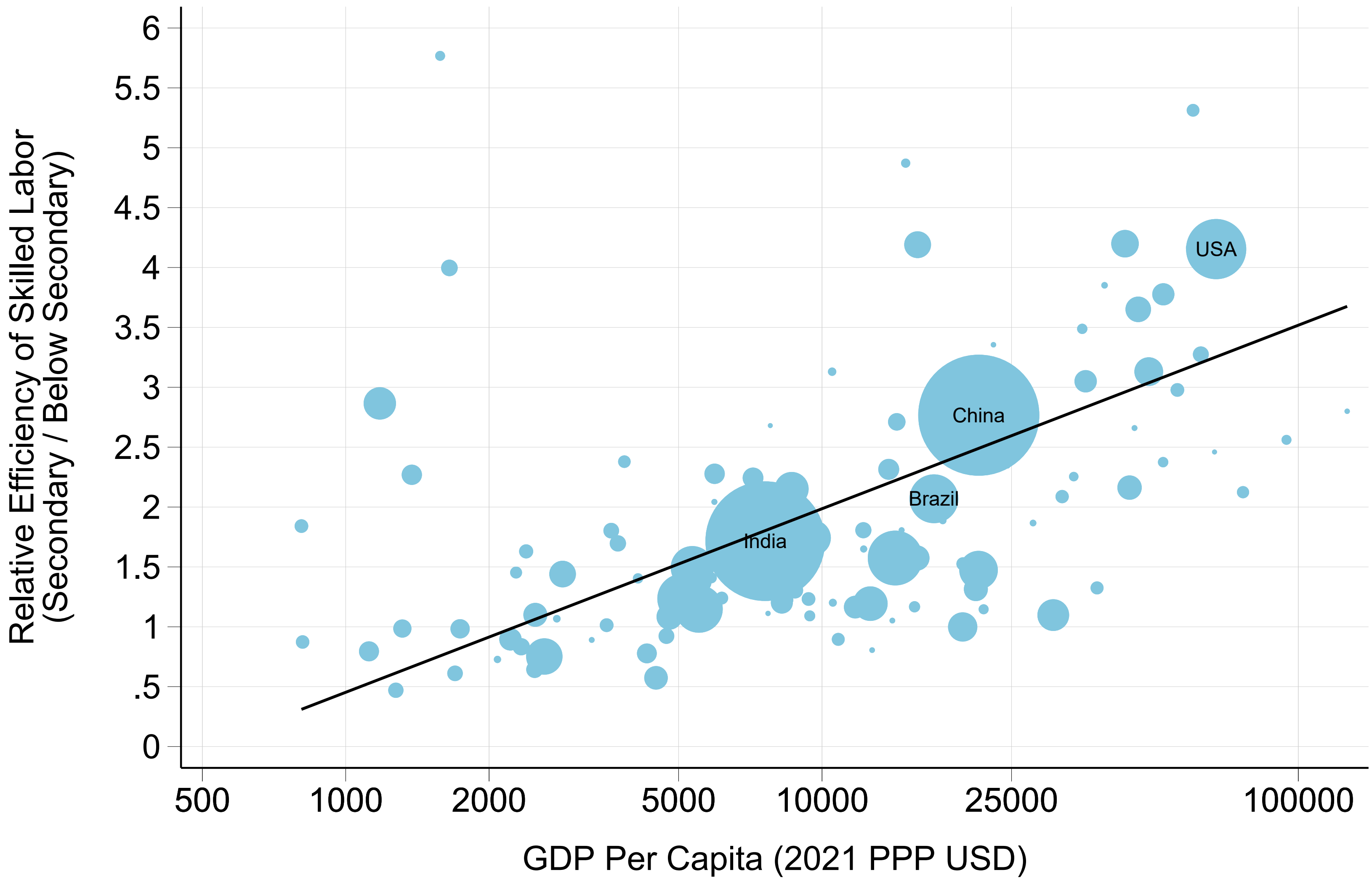
Notes. The figure compares the 2019 (final) and true total return to tertiary education (the percent increase in personal income of moving from secondary to tertiary education) in each country. True returns are estimated assuming an elasticity of substitution of 5.

Figure B6 – Relative Efficiency of Skilled Labor Versus GDP Per Capita (Tertiary Versus Below Tertiary)



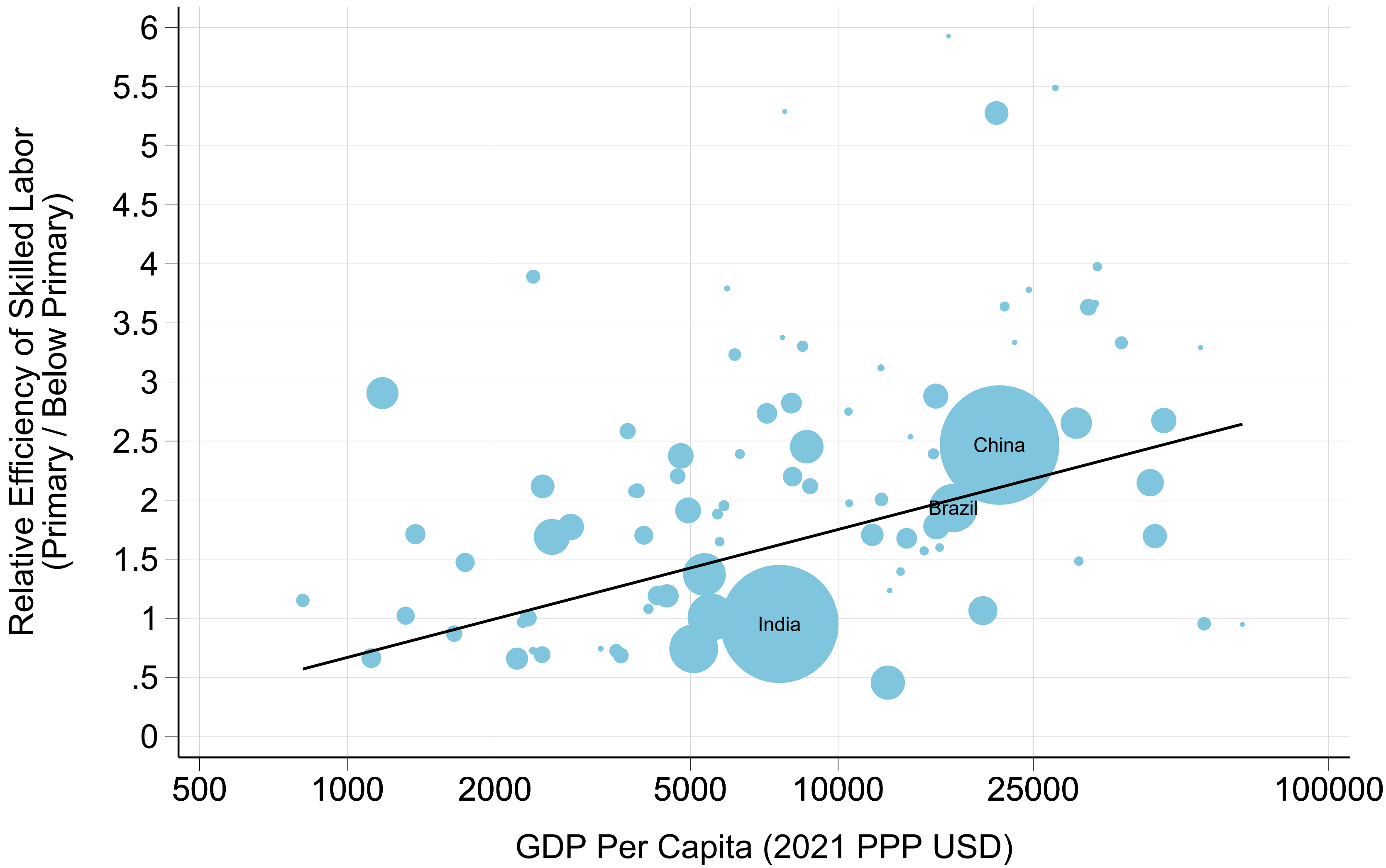
Notes. Author's calculations using survey microdata.

Figure B7 – Relative Efficiency of Skilled Labor Versus GDP Per Capita (Secondary Versus Below Secondary)



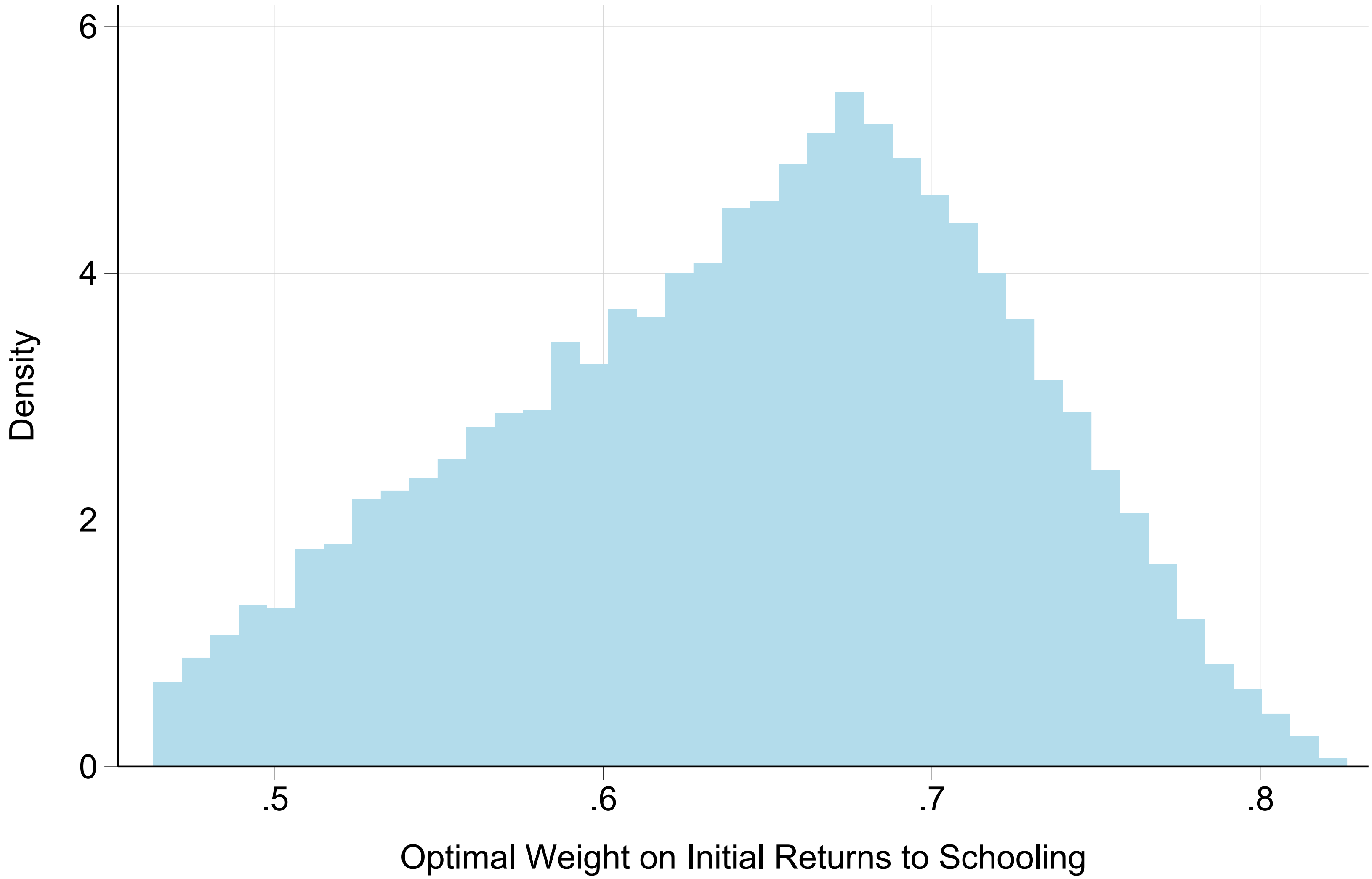
Notes. Author's calculations using survey microdata.

Figure B8 – Relative Efficiency of Skilled Labor Versus GDP Per Capita (Primary Versus No Schooling)



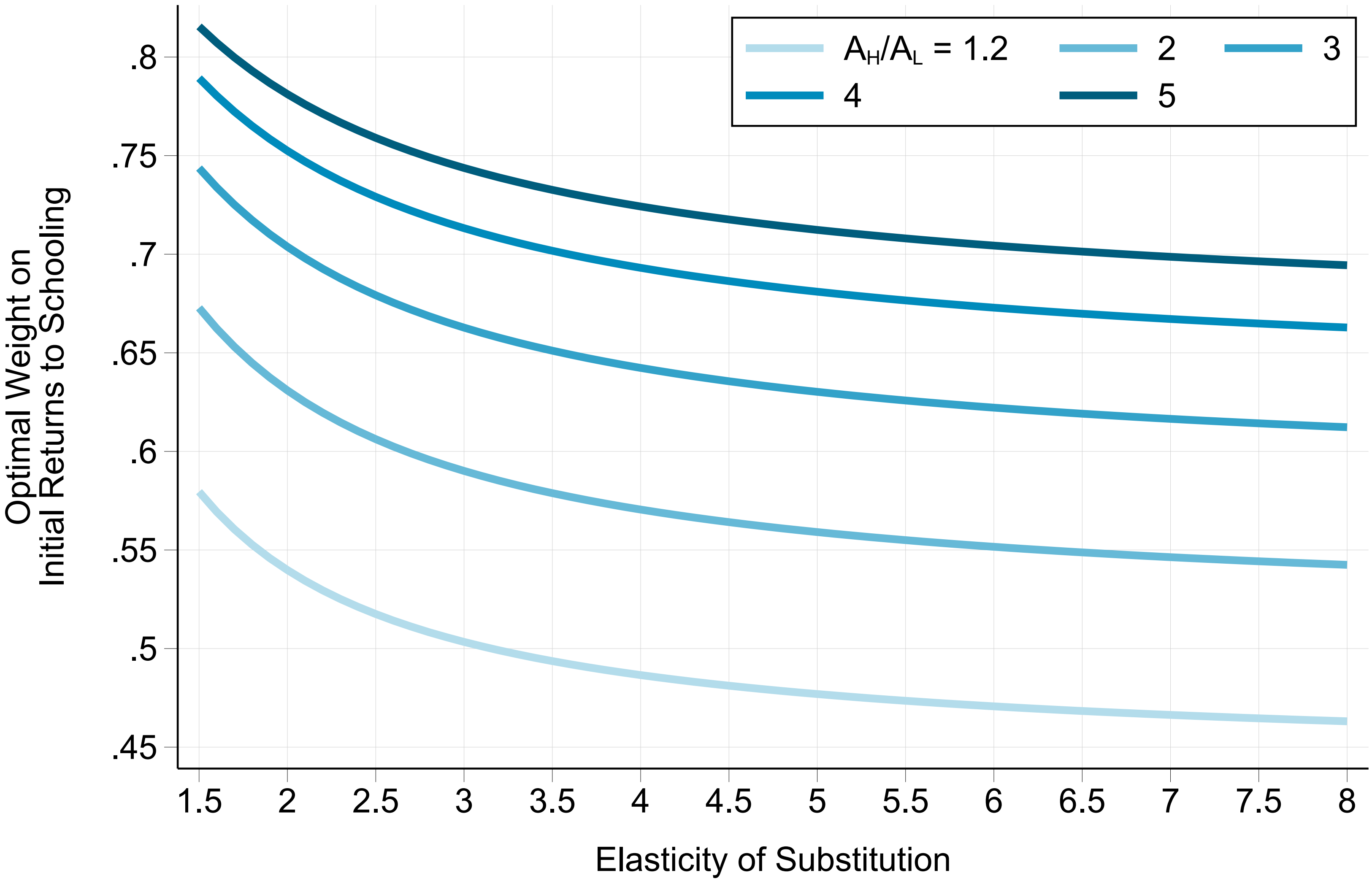
Notes. Author's calculations using survey microdata.

Figure B9 – Returns to Schooling: Simulated Distribution of Optimal Weights on Initial Returns



Notes. The figure plots the distribution of optimal weights on initial returns to schooling required to estimate the true effect of educational expansion on output, based on simulations varying the elasticity of substitution, the relative efficiency of skilled workers, and the magnitude of educational expansion.

Figure B10 – Returns to Schooling: Optimal Weights on Initial Returns Under Different Parametrizations of the CES Production Function



Notes. The figure plots how optimal weights on initial returns to schooling required to estimate the true effect of educational expansion on output vary, depending on parametrizations of relative skill efficiency A_H/A_L and the elasticity of substitution σ.

Table B1 – Empirical Estimates of the Elasticity of Substitution Between Skill Groups

Source	Country	Tertiary/Below	Secondary/Below
Long-run elasticity			
Bils, Kaymak, and Wu (2022)	Cross-Country	4 to 6	4 to 6
Hendricks and Schoellman (2023)	Cross-Country	4.5	7.8
Short-run elasticity			
Bowlus et al. (2021)	United States	5.3	
Autor, Goldin, and Katz (2020)	United States	1.62	
Hershbein, Kearney, and Pardue (2020)	United States	1.4	
Acemoglu and Autor (2011)	United States	1.6	
Goldin and Katz (2007)	United States	1.6	2 to 5
Ciccone and Peri (2005)	United States	1.5	
Heckman, Lochner, and Taber (1998)	United States	1.4	
Katz and Murphy (1992)	United States	1.41	
Murphy, Riddell, and Romer (1998)	Canada	1.36	
Angrist (1995)	Palestine	2	
Vu and Vu-Thanh (2022)	Vietnam	2.67	
Fernández and Messina (2018)	Latin America	1.25	2.3
Khanna (2023)	India		4.24
Caselli and Coleman (2006)	Cross-Country	1.3	

Notes. The table reports selected estimates of the elasticity of substitution between skill groups from various empirical studies. Bils, Kaymak, and Wu (2022): unique elasticity of substitution for all skill groups.

C. Natural Experiments

This appendix exploits evidence from three natural experiments to shed light on the ability of the model to reproduce results from real-world episodes of educational expansion. Section C.1 outlines the general econometric framework. Sections C.2, C.3, and C.4 turn to analyzing the Indian District Primary Education Program, the Indonesian INPRES school construction program, and U.S. compulsory schooling laws. Overall, the model does a remarkable job at reproducing aggregate and distributional effects of actual policies. If anything, it tends to underestimate the effect of human capital accumulation on real earnings at the bottom of the income distribution. Estimates relying on the simulation method outlined in section 2.4 should thus be considered a lower bound.

C.1. General Methodology

A large literature focuses on causally identifying individual returns to schooling. Less is known of the distributional effects of increasing human capital at the level of regions or countries. This section attempts to shed some light on these effects by studying three large-scale natural experiments in India, Indonesia, and the United States. More specifically, consider the following empirical specification:

$$\ln y_{rt}^{i} = \gamma_0^{i} + \gamma_1^{i} S_{rt} + X_{rt}^{i}\beta + \delta_r + \delta_t + \varepsilon_{rt} \tag{48}$$

$$S_{rt} = \alpha_0 + \alpha_1 Z_{rt} + \eta_{rt} \tag{49}$$

Where i denotes income groups, such as quintiles, in subnational regions r at time t. The objective is to estimate the impact of increasing average regional schooling S_{rt} on y_{rt}^{i}, the log average income of income group i. X_{rt}^{i} is a vector of controls, such as the demographic composition of the region, δ_r are subnational region fixed effects, and δ_t are time fixed effects. The parameter of interest is γ_1^{i}, the semi-elasticity of average income of group i to regional average years of schooling. One option is to directly estimate equation 48 by OLS. This is analogous to the usual cross-country or cross-region growth regression specification (e.g., Gennaioli et al., 2013). Alternatively, average schooling S_{rt} can be instrumented using an instrument Z_{rt}, such as compulsory schooling laws, which generates quasi-random differential trends in average schooling across regions. This approach has also been used, in particular in the case of U.S. compulsory schooling laws, mainly with the objective of estimating human capital externalities (Acemoglu and Angrist, 2000; Ciccone and Peri, 2006; Guo, Roys, and

Seshadri, 2018). The main addition here is the focus on distributional effects, which amounts to estimating heterogeneous treatment effects by income group.

Estimating the distributional effects of educational expansion is empirically challenging, because it requires two sets of data that are rarely jointly available: data on the distribution of income within subnational regions, and an instrument that can predict quasi-random variation in regional schooling. Drawing on existing work, I study three such sources of variation: the India District Primary Education Program, the Indonesian School Construction Program, and U.S. state compulsory schooling laws.

C.2. India District Primary Education Program, 1994-2004

C.2.1. Context

Between the 1990s and the beginning of the 2000s, India engaged in a massive expansion of public schooling, the District Primary Education Program (DPEP), targeting low-literacy regions. Districts with a female literacy rate below the national average were more likely to benefit from the policy. Exploiting this allocation rule, Khanna (2023) estimates the general equilibrium effects of the program using a regression discontinuity design. He finds a return to schooling of about 13% per year (after accounting for general equilibrium effects). General equilibrium effects induced by the greater relative supply of skilled workers depress returns by one-third, while indirectly benefiting unskilled workers. To the best of my knowledge, this represents one of the only studies providing quasi-experimental evidence on the aggregate effects of schooling expansion initiatives. The design and data make it particularly well-suited for estimating the distributional incidence of human capital accumulation.

C.2.2. Data

I exploit data from the replication package provided by Khanna (2023). Exposure to the program is determined by district female literacy in 1991. There are 571 districts, 271 of which were treated by the program. Individual outcomes are obtained from the 2009 National Sample Survey (NSS). The microdata covers information on wages and education at the district level, allowing for a direct estimation of the impact of the program on the distribution of labor income. As in Khanna (2023), the sample is restricted to all adults aged 17 to 75 with positive wage income.

C.2.3. Empirical Specification

I follow Khanna (2023) and estimate the impact of the policy using the same regression discontinuity design as in the paper, comparing districts below and above the average female literacy rate. Optimal bandwidths are calculated using either the Calonico, Cattaneo, and Titiunik (2014) method or the Imbens and Kalyanaraman (2012) method (henceforth CCT and I and K, respectively).

The main addition is that I focus on the effect of the program on the average wage of each wage quintile, to directly get a reduced-form estimate of the distributional incidence of primary education expansion. Khanna (2023) centers his analysis on the estimation of individual returns to schooling, as well as spillovers to other skill groups. In contrast, I use the RD to directly instrument average district schooling and estimate its impact on the average wage of each wage quintile within each district.

Figure C1 plots the first stage, comparing district average years of schooling among adults with positive wage income below and above the literacy cutoff. Districts below the cutoff were more likely to benefit from the program. Adults living in districts that were just below this cutoff have significantly higher levels of education than those living in districts just above.

C.2.4. Results

Table C1 presents the main results. Increasing district average years of schooling by one year is associated with a 12% increase in wages in treated districts (CCT method). This effect is almost two times larger for the bottom 20% of earners, who benefit from a 21% increase in wages. In contrast, the top 20% see their average wage decline, although the coefficient is not statistically significant. Results relying on the I and K method are similar, but the aggregate effect of educational expansion appears even larger, reaching 26% for average wages and 32% for the bottom 20%. Aggregate returns to schooling estimated using this method are in the range of individual returns estimated by Khanna (2023), who finds returns of 16% (CCT) to 21% (I and K) using conventional 2SLS estimates, and 13% (CCT) after accounting for general equilibrium effects.

Table C2 compares the CCT estimates to simulated effects of expanding primary education, under different parametrizations of the return to schooling and the elasticity of substitution between skilled and unskilled workers. Figure 6 plots the corresponding coefficients in the specific case where the return to schooling is set to 13% and the elasticity of substitution to 4 (corresponding approximately to the values obtained by Khanna, 2023). The simulation is done by upgrading the education of randomly sampled individuals from no schooling to primary education, increasing their earnings using the return to schooling, and finally adjusting relative wages for general equilibrium effects, using the method outlined in section 2.

Simulated estimates fall close to the true effects of the policy. Simulation results show that with a return to schooling of 13%, increasing average district education by one year through basic education is associated with an increase in average wages of about 9%. The aggregate effect is lower than the individual return, because those benefiting from the expansion are workers with no schooling, whose wages are significantly lower than average. To simulate an aggregate effect similar to the one estimated using the natural experiment, a higher individual return to schooling is required, in the order of 16%. Relying on individual returns to estimate the effect of basic education expansion thus provides a lower bound on the true aggregate effect of educational expansion.

The simulation also predicts distributional effects that are very similar to those estimated with the RD design. Both in the simulation and in the natural experiment, benefits appear relatively similar for the first four quintiles and significantly lower for the top 20%. This can be rationalized by the fact that in India, workers with no schooling and workers with basic education are both prevalent among the bottom 80% of the distribution, so that upgrading some workers from no schooling to basic education benefits this entire group. Simulated effects do not vary much with the elasticity of substitution, although lower values of the elasticity are associated with greater gains for the bottom quintile, where the concentration of workers with no schooling in the greatest. All in all, the model performs remarkably well at reproducing the observed economic effects of primary education expansion in India.

C.3. Indonesia School Construction Program, 1973-1978

C.3.1. Context

Between 1973 and 1978, Indonesia engaged in a massive school construction program aiming to expand access to basic education throughout the country. Exploiting differences in exposure to newly built schools across cohorts and regions, Duflo (2001) estimates individual returns to schooling ranging from 7% to 11%. A number of studies have updated and extended her analysis since then, focusing on intergenerational effects (Akresh, Halim, and Kleemans, 2023), structural transformation (Karachiwalla and Palloni, 2019), or rural-urban migration (Hsiao, 2023).

Duflo (2004) also moves beyond individual outcomes to focus on spillovers of the program to non-treated groups. Combining labor force surveys covering the 1986-1999 period, she estimates the impact of the greater supply of young skilled workers on older generations' formal employment and wages. Her analysis shows mixed findings, suggesting a decline in the wages of non-treated groups, but an increase in employment in the formal sector.

C.3.2. Data

Drawing on the work of Duflo (2004), I exploit differential exposure to the program by district to estimate the aggregate and distributional effects of primary education expansion. My analysis expands her work in two ways. First, I significantly expand the time coverage of the data, which increases statistical power and allows me to get closer to long-run effects. To do so, I collect and harmonize every round of the SUSENAS, a household survey covering about a million individuals every year, from 1993 to 2019. The result is a balanced panel of 230 districts, for which I have yearly data on the education level of the labor force, the distribution of consumption, and other sociodemographic variables over a twenty-six-year period.[28] Second, I study the effects of the program on total district consumption and its distribution by consumption quintile, while Duflo (2004) focuses on spillover effects on older cohorts. The sample is restricted to all adults aged 15 to 70; consumption is then split equally all members of the household.

C.3.3. Empirical Specification

The empirical specification corresponds to the one in equation 48, with schooling being instrumented as in Duflo (2004). I estimate the effect of average years of schooling in district r on the log average consumption of decile i, controlling for district and year fixed effects. Average years of schooling is instrumented by the interaction between survey years and the number of schools built per 5-14 population between 1974 and 1978.[29] The school construction program is thus taken as an instrument for *differential trends* in the education of the working-age population across districts from 1993 to 2019. Districts with greater treatment intensity are expected to see a faster secular increase in average schooling, because of the greater access to schooling enabled by the policy for cohorts educated 20-50 years ago. The identification assumption, analogous to Duflo (2004), is that there is no unobserved shock both correlated with the program and affecting household expenditure during that period.

[28]Some districts have undergone splits and merges over the period of interest. I rely on crosswalks provided by Roodman (2022) to ensure consistent boundaries over time.

[29]Given significant noise introduced by the low sample size available for each district-year cell, I specify survey years as a continuous variable in the first stage. Indeed, as shown by Duflo (2004), we should expect the program to have introduced smooth, secular differential trends in educational expansion. Constraining the interaction of survey years and treatment intensity to follow such a secular trend makes the results less sensitive to different empirical strategies.

Figure C2 plots the first stage. The dependent variable is average years of schooling in a given district-year; each point corresponds to the coefficient of the interaction of a survey year dummy with treatment intensity, with 1993 taken as the baseline year. This figure is analogous to Duflo (2004), figure 2. In districts with greater exposure to the school construction program in 1974-1978, average years of schooling among the working-age population have risen at a significantly faster pace. The estimates are slightly noisy, because of the relatively low number of observations available in each district-year cell, but they confirm that the program had long-lasting effects on regional rates of human capital accumulation.

C.3.4. Results

Table C3 presents the main results. The baseline specification controls for the demographic and gender composition of each district, the share of college graduates, and district and year fixed effects. Increasing average district schooling by one year is associated with an 8.7% rise in average consumption in the district. This effect is almost four times larger for the bottom quintile (22%) than for the top quintile (5.8%). Columns 4 to 6 add controls for 1971 primary school enrollment and water and sanitation spending interacted with survey year, as in Duflo (2004). These estimates are underpowered and the standard errors much larger, because of limited sample size in each district, but the results obtained are qualitatively similar. The effect on bottom 20% average consumption rises to 51%, while that on the top 20% boils down to zero. Columns 7 to 9 add further controls for 1971 child population and population density interacted with survey year. This model is even more underpowered, but the point estimates remain of the same order of magnitude. In particular, the coefficient on the average income of the bottom 20% remains large (45%) and statistically significant at the 5% level. While it is clear that the sample size is not sufficient to precisely estimate aggregate returns to schooling, the progressive nature of the policy stands out across all specifications.

Table C2 compares the benchmark estimates to simulated effects of the policy using the 1996 Indonesian labor force survey (SAKERNAS). Figure 7 plots the coefficients by quintile when the return to schooling is set to 12% and the elasticity of substitution to 4.[30] The simulation is done exactly as in the Indian case, upgrading the education of randomly sampled individuals from no schooling to primary education, increasing their earnings using the return to schooling, and finally adjusting relative wages.

[30]Duflo (2004) finds no evidence that schooling expansion led to a significant decline in the skill premium (Duflo, 2004, table 6), which would point to an infinite elasticity. However, the standard errors are very large, implying confidence intervals that include both very low and negative elasticities. Rather than evidence in favor of perfect substitution, these findings point to the fact that limitations in the sample size unfortunately make it difficult to estimate such elasticity with the available data.

As in India, the simulation does a good job at reproducing results from the natural experiment. The expansion of primary education is estimated to be progressive in all specifications, with orders of magnitude similar to those found in the data. Lower values for the elasticity of substitution are associated with significantly higher growth for the bottom 40% relative to the top 60%. The benchmark specification, with a return of 11% (close to the estimate of Duflo, 2001) and an elasticity of 4, matches both aggregate and distributional effects particularly well. Higher elasticities of substitution generally imply inequality-reducing effects of the policy that are too low in comparison to those observed in the data.

C.4. U.S. Compulsory Schooling Laws, 1875-1961

C.4.1. Context

Between the mid-19th and the mid-20th century, U.S. states gradually implemented laws limiting child labor and enforcing compulsory school attendance for newly educated cohorts. The effect of these laws were first studied by Acemoglu and Angrist (2000), who combined data on laws implemented from 1914 to 1965 with census microdata to estimate the magnitude of human capital spillovers. Their analysis gave rise to a rich literature exploiting compulsory schooling laws to estimate individual returns to schooling (Clay, Lingwall, and Stephens, 2021; Stephens and Yang, 2014), elasticities of substitution between skill groups (Ciccone and Peri, 2006), and human capital externalities (Ciccone and Peri, 2006; Guo, Roys, and Seshadri, 2018; Iranzo and Peri, 2009).

C.4.2. Data

My analysis relies on similar sources than those used in the existing literature, but extends previous work in two ways. First, I study the total aggregate and distributional effects of educational expansion, while existing studies focus on estimating different dimensions of these effects separately. Second, I exploit recently compiled data by Clay, Lingwall, and Stephens (2021), covering compulsory schooling laws over the entire 1875-1961 period. This represents an important improvement over the previous literature, which only covered laws implemented after 1915, based on the database of Acemoglu and Angrist (2000).[31] To estimate the impact of schooling on the distribution of income, I rely on the 1940 to 2000 census microdata samples

[31]Consider in particular the 1940 to 1960 censuses, which cover periods of the twentieth century during which basic education mattered most for explaining cross-state variations in human capital. Post-1915 compulsory schooling laws fail to capture variations in schooling for all workers older than 25-45 during that period, so they end up missing important sources of variations.

available from IPUMS USA, which cover personal income, state of birth, state of residence, education, and other sociodemographic variables by ten-year interval. The sample is restricted to all adults aged 25 to 65 with positive personal income (wage income in 1940) living in the contiguous United States.

C.4.3. Empirical Specification

As in the Indian and Indonesian case studies, I regress the average income of each personal income decile on average state schooling, instrumented by compulsory schooling laws. More specifically, consider the following instrument for average years of schooling S_{st} in state s at time t:

$$S_{st} = \pi_0 + \pi_1 \sum_c \sum_{s'} N_{css't} RS_{cs'} + \theta_s + \theta_t + u_{st} \tag{50}$$

Where $RS_{cs'}$ is required years of schooling for cohort c born in state s', $N_{css't}$ is the number of individuals living in state s at time t who were born in state s', and θ_s and θ_t are state and year fixed effects. Required years of schooling correspond to the time a children born in a given year is required to stay in school, calculated by combining information on required attendance at each year of life (see Clay, Lingwall, and Stephens, 2021; Stephens and Yang, 2014). The instrument is thus equal to the average *required* years of schooling of the working-age population of state s, calculated by averaging required years of schooling across all cohort-state-of-birth cells, weighted by their relative populations at time t. This approach is analogous to the one recently adopted by Guo, Roys, and Seshadri (2018), who instrument average state education by required years of schooling in each state-age cell.

The interpretation of the instrumentation strategy is similar to that of the Indonesian case study. Differences in required years of schooling across cohorts born from 1875 to 1961 are used to predict differential trends in average schooling across states from 1940 to 2000.

Figure C3 provides a concrete illustration of how required years of schooling evolved across cohorts born in Alabama, California, Indiana, and Massachusetts from 1875 to 1965. In 1875, Massachusetts was the only state imposing compulsory education, for a duration of 6 years. All states saw the implementation of increasingly restrictive laws, but with significant variations in timing and intensity. Indiana rapidly shifted from no compulsory schooling to nine years in 1899, while Alabama followed the same transition much more gradually, from no compulsory law until 1902 to four years in 1905, six years in 1909, eight years in 1919, and finally nine years by 1933.

Table C5 shows the first stage. Column 1 controls for the demographic, gender, and racial composition of each state, as well as the share of college graduates. An additional average required year of schooling is associated with a 0.19 increase in actual average years of schooling among the working-age population. Column 2 add census × year fixed effects, which have been shown to potentially matter significantly when estimating the effect of U.S. compulsory schooling laws (Stephens and Yang, 2014). This reduces the effect to 0.14. Finally, column 3 adds further controls for initial conditions, interacting census year fixed effects with average income and average years of schooling in 1940. This is a very ambitious specification, as it implies estimating over 100 coefficients on a sample of only 343 observations. The coefficient on required years of schooling is reduced to 0.12, and remains statistically significant at the 1% level.

C.4.4. Results

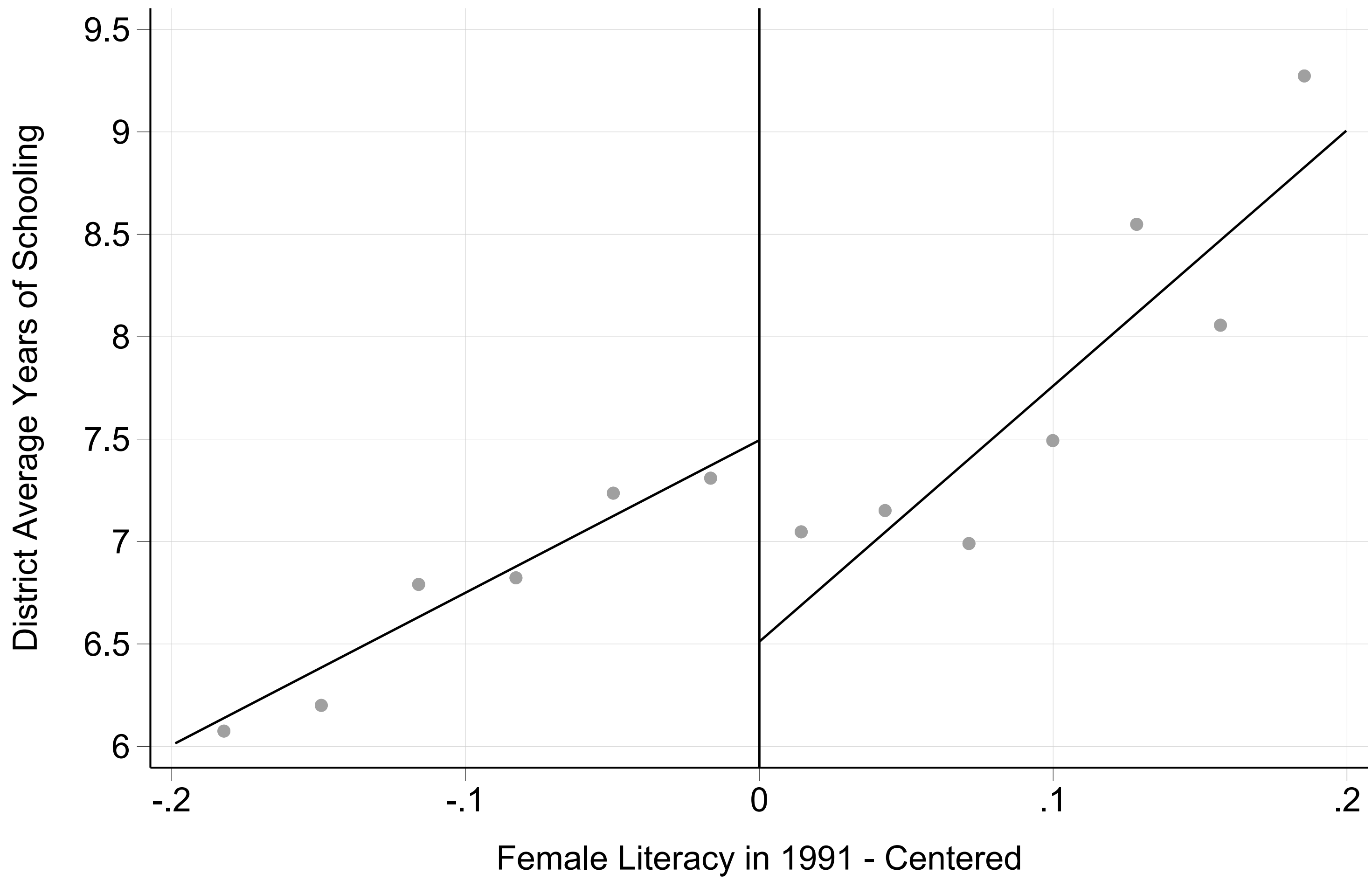
Table C6 presents the main results. In the baseline specification, an additional average year of schooling is associated with a 0.16 log-point increase in average income. The corresponding values are 0.44 for the bottom 20%, compared to 0.05 for the top 20%. Educational expansion thus appears to have been a powerful driver of inequality reduction, even more so in the U.S. than in India and Indonesia. Adding interacted census region and year fixed effects leaves the results almost unchanged (columns 4 to 6). Columns 7 to 9 further add controls for initial conditions. The aggregate effect is slightly lower (0.08), and the estimates are unsurprisingly underpowered. Even under this highly demanding specification, however, the coefficient on the bottom 20% remains large (0.27) and statistically significant at the 5% level.

Table C7 compares observed and simulated effects of the policy. Figure 8 plots the coefficients by quintile when the return to schooling is set to 12% and the elasticity of substitution to 4, as in the previous case studies. The simulation is done by upgrading the education of randomly sampled individuals with either no schooling or primary education to secondary education, given that required years of schooling range from 0 to 9 years.

Here, in contrast to the two previous case studies, the model appears to strongly underestimate the aggregate and inequality-reducing effects of the policy. Even with returns to schooling of 16% and an elasticity of substitution of 2, it can generate an effect on the average income of the bottom quintile of "only" 0.4 log points, while overestimating growth for the top quintile.

There are at least three reasons why this might be the case. First, state compulsory schooling laws extended both primary and secondary school attendance, with significant variations in timing and intensity across states. This makes it more difficult to accurately simulate the overall effect of these policies. Indeed, who exactly benefited from them (individuals who would have had either no schooling, some primary education, or some secondary education in the absence of these policies) is less clear than in the Indian and Indonesian cases. Second, there is evidence that returns to schooling were substantially higher at the bottom of the income distribution during the first wave of compulsory schooling laws (Clay, Lingwall, and Stephens, 2021). In contrast, the simulation assumes a constant return by income group. This limits by construction its ability to capture higher returns for low-income earners. Third, recent evidence points to potentially large human capital externalities from schooling expansion in the United States, as high as 6-8% per year of schooling (Guo, Roys, and Seshadri, 2018). This might explain why even with an individual return to schooling as high as 16%, the simulation ends up underestimating the aggregate return by 4-5 percentage points.

Figure C1 – India DPEP: First Stage



Notes. The figure compares average years of schooling among adult wage earners below and above the literacy cutoff used to allocate the program. Data from Khanna (2023).

Table C1 – India DPEP: Aggregate and Distributional Effects of Schooling

	Bandwidth Selection: CCT Method			Bandwidth Selection: I and K Method		
	Average Income	Bottom 20% Income	Top 20% Income	Average Income	Bottom 20% Income	Top 20% Income
Average Years of Schooling	0.117*	0.207***	-0.092	0.257***	0.316***	0.012
	(0.061)	(0.059)	(0.071)	(0.058)	(0.058)	(0.068)
N	46314	9007	9515	46314	9007	9515

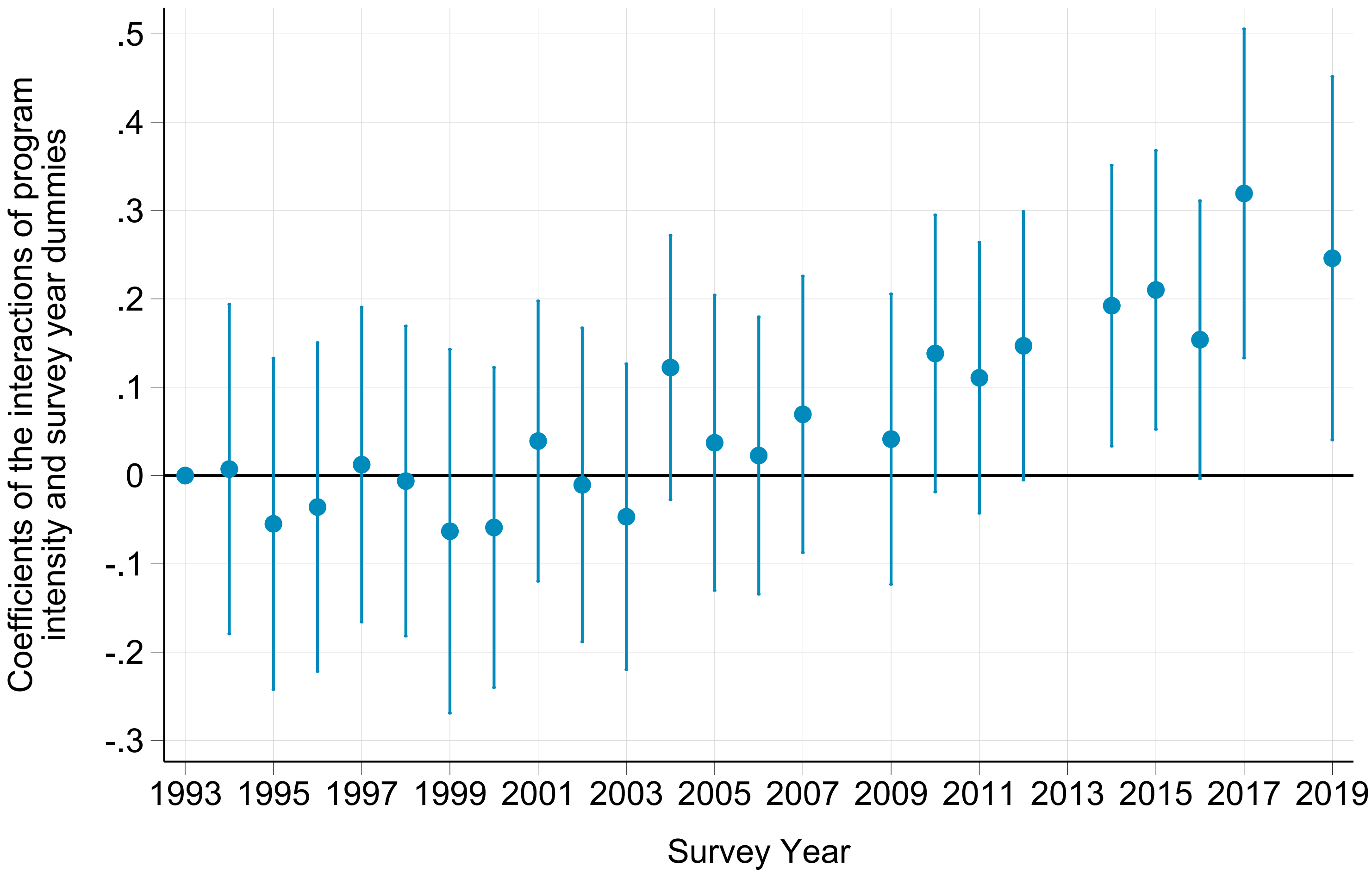
Notes. The table reports the effect of district average years of schooling on district average income, the average of the bottom 20%, and the average income of the top 20%. Bandwidths: "CCT" indicates the Calonico, Cattaneo, and Titiunik (2014) method, "I and K" the Imbens and Kalyanaraman (2012) method. Data from Khanna (2023). Standard errors in parenthesis. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C2 – India DPEP: Actual vs. Simulated Effects of Educational Expansion

	Parameters		Effect of Increasing Average District Schooling by One Year (%)					
	Return to Schooling	Elasticity of Substitution	Average Income	Q1	Q2	Q3	Q4	Q5
Actual Effect			11.7	20.7	11.2	17.1	19.3	-9.2
Simulated Effect	13%	∞	9.3	17.4	13.5	14.8	10.0	5.6
	13%	6	9.3	19.2	13.4	14.8	10.0	5.5
	13%	4	9.3	20.0	13.3	14.8	10.0	5.4
	13%	2	9.3	22.6	12.9	14.9	9.9	5.2
	16%	∞	12.4	18.9	15.0	18.5	14.3	8.3
	16%	6	12.4	20.9	15.0	18.6	14.3	8.1
	16%	4	12.4	21.9	15.0	18.6	14.3	8.0
	16%	2	12.4	24.8	14.7	18.9	14.2	7.7
	20%	∞	17.1	20.0	16.5	21.9	21.0	13.4
	20%	6	17.1	22.3	16.8	22.2	21.0	13.1
	20%	4	17.1	23.4	16.9	22.4	20.9	12.9
	20%	2	17.1	26.8	17.1	22.9	20.8	12.4

Notes. Actual effect: estimated effect of the policy on average district income and the average income of each wage quintile, using data from Khanna (2023). Simulated effect: effect of the policy predicted using 2019 LFS data, under different assumptions on the return to a year of schooling and the elasticity of substitution between skilled and unskilled workers.

Figure C2 – Indonesia INPRES: First Stage: Effect of the Program on District Average Years of Schooling, 1993-2019



Notes. The figure compares the evolution of average years of schooling in districts with more or less exposure to the INPRES school construction program. The dependent variable is average years of schooling in each district-year. Estimates combine 1993-2019 SUSENAS microdata with treatment intensity by district from Duflo (2001).

Table C3 – Indonesia INPRES: Aggregate and Distributional Effects of Schooling

	Baseline			+ Controlling for 1971 Primary School Enrollment and Water & Sanitation Spending			+ Controlling for 1971 Child Population and Population Density		
	Average Income	Bottom 20% Income	Top 20% Income	Average Income	Bottom 20% Income	Top 20% Income	Average Income	Bottom 20% Income	Top 20% Income
Average Years of Schooling	0.087***	0.220***	0.058*	0.133	0.505***	-0.002	0.084	0.445**	-0.029
	(0.026)	(0.035)	(0.031)	(0.083)	(0.145)	(0.097)	(0.108)	(0.179)	(0.131)
N	5520	5520	5520	5352	5352	5352	5304	5304	5304

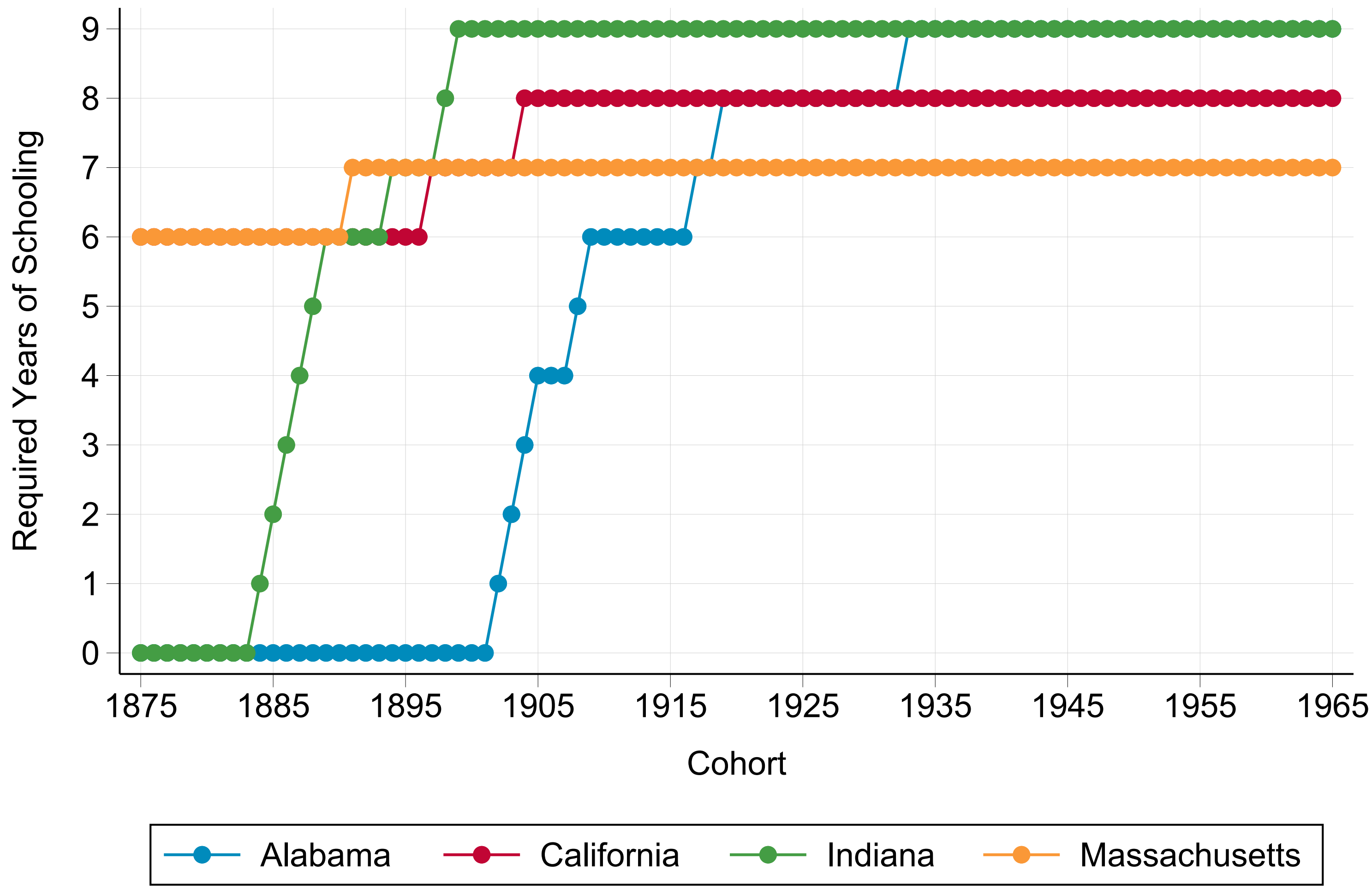
Notes. The table reports the effect of regency average years of schooling on regency average income, the average income of the bottom 20%, and the average income of the top 20%. Columns 1 to 3 control for the demographic composition of the regency, the share of women, and the share of workers with tertiary education. Columns 4 to 6 add controls for 1971 primary school enrollment rates and water and sanitation spending, interacted with survey year. Columns 7 to 9 further add controls for the share of the population aged 5 or below in 1971 and population density in 1971, interacted with survey year. Data from Duflo (2001) and Roodman (2022). Standard errors in parenthesis. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C4 – Indonesia INPRES: Actual vs. Simulated Effects of Educational Expansion

	Parameters		Effect of Increasing Average District Schooling by One Year (%)					
	Return to Schooling	Elasticity of Substitution	Average Income	Q1	Q2	Q3	Q4	Q5
Actual Effect			8.7	22.0	15.4	11.5	8.1	5.8
Simulated Effect	9%	∞	5.7	15.4	10.6	5.3	4.6	4.0
	9%	6	5.7	19.1	11.6	6.4	5.2	3.7
	9%	4	5.7	20.4	12.3	7.1	5.5	3.5
	9%	2	5.7	25.4	13.9	4.4	3.2	3.0
	11%	∞	7.7	17.2	12.9	7.6	7.3	5.7
	11%	6	7.7	20.6	13.8	7.5	7.1	5.4
	11%	4	7.7	22.5	14.6	8.6	7.8	5.3
	11%	2	7.7	27.9	16.4	6.0	5.8	5.1
	13%	∞	10.5	18.8	15.1	9.6	10.0	8.9
	13%	6	10.5	22.0	15.8	9.3	9.8	8.6
	13%	4	10.5	23.9	16.2	8.9	9.7	8.6
	13%	2	10.5	28.8	18.6	7.6	8.2	8.2

Notes. Actual effect: estimated effect of the policy on average district income and the average income of each wage quintile, using data from Duflo (2001). Simulated effect: effect of the policy predicted using 1996 SAKERNAS microdata, under different assumptions on the return to a year of schooling and the elasticity of substitution between skilled and unskilled workers.

Figure C3 – U.S. Compulsory Schooling Laws: Examples



Notes. Author's elaboration based on data from Clay, Lingwall, and Stephens (2021).

Table C5 – U.S. Compulsory Schooling Laws: First Stage
Effect of Required Years of Schooling on State Average Years of Schooling

	(1)	(2)	(3)
Required Years of Schooling	0.191***	0.141***	0.115***
	(0.032)	(0.038)	(0.032)
Region × Year FE	No	Yes	Yes
Extended Controls	No	No	Yes
N	343	343	343

Notes. The unit of observation is the state-year. Required years of schooling: average required years of schooling in each state-year, instrumented using required years of schooling for each state-cohort. Region × Year FE: interacted census region and census year fixed effects. Extended controls: additional controls for 1940 average years of schooling and average personal income interacted with census year fixed effects.

Table C6 – U.S. Compulsory Schooling Laws: Aggregate and Distributional Effects of Schooling

	Baseline			+ Census Region × Year FE			+ Controls for 1940 Educational Attainment and Average Income × Year FE		
	Average Income	Bottom 20% Income	Top 20% Income	Average Income	Bottom 20% Income	Top 20% Income	Average Income	Bottom 20% Income	Top 20% Income
Average Years of Schooling	0.157***	0.437***	0.050*	0.147***	0.431***	0.079*	0.082	0.272**	0.063
	(0.032)	(0.095)	(0.027)	(0.045)	(0.110)	(0.044)	(0.051)	(0.114)	(0.057)
N	343	343	343	343	343	343	343	343	343

Notes. The table reports the effect of state average years of schooling on state average income, the average income of the bottom 20%, and the average income of the top 20%. Columns 1 to 3 control for the demographic, gender, and racial composition of each state, as well as the share of workers with tertiary education. Columns 4 to 6 add census region × year fixed effects. Columns 7 to 9 further add controls for 1940 average years of schooling and average personal income, interacted with survey year dummies. Data from IPUMS census microdata combined with information on compulsory schooling laws from Acemoglu and Angrist (2000) and Clay, Lingwall, and Stephens (2021). Standard errors in parenthesis. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table C7 – U.S. Compulsory Schooling Laws: Actual vs. Simulated Effects of Educational Expansion

	Parameters		Effect of Increasing Average State Schooling by One Year (%)					
	Return to Schooling	Elasticity of Substitution	Average Income	Q1	Q2	Q3	Q4	Q5
Actual Effect			15.7	43.7	46.0	25.9	12.3	5.0
Simulated Effect	8%	∞	4.5	17.4	9.6	5.2	3.6	2.8
	8%	6	4.5	20.9	11.4	5.7	3.3	2.1
	8%	4	4.5	22.6	12.3	6.0	3.2	1.8
	8%	2	4.5	27.8	14.8	6.8	2.8	0.9
	12%	∞	7.4	24.0	14.0	7.6	5.5	5.9
	12%	6	7.4	27.5	15.7	8.2	5.3	5.3
	12%	4	7.4	29.2	16.6	8.4	5.2	4.9
	12%	2	7.4	34.4	19.1	9.2	4.7	4.0
	16%	∞	11.0	29.2	18.3	9.7	7.4	10.6
	16%	6	11.0	32.7	20.1	10.3	7.2	9.9
	16%	4	11.0	34.5	20.9	10.6	7.1	9.5
	16%	2	11.0	39.7	23.4	11.4	6.7	8.6

Notes. Actual effect: estimated effect of the policy on average state income and the average income of each personal income quintile, combining IPUMS census microdata with information on compulsory schooling laws from Acemoglu and Angrist (2000) and Clay, Lingwall, and Stephens (2021). Simulated effect: effect of the policy predicted using 1960 census microdata, under different assumptions on the return to a year of schooling and the elasticity of substitution between skilled and unskilled workers.

D. Education Quality

A natural question is how education quality might have changed from 1980 to 2019, and potential implications for the results presented in this paper. The main source of concern is that if education quality has increased or decreased, then educational attainment becomes a biased measure of actual changes in the education of the labor force. If quality has changed, 1980 and 2019 levels of attainment are not comparable indicators anymore. To make them comparable, one would need to adjust them for changes in quality by, for instance, re-expressing 2019 years of schooling in 1980 quality-adjusted equivalents.

This section discusses existing evidence on the evolution of education quality in developed and developing economies, and attempts to quantify how sensitive are the main results to accounting for these changes. Existing sources provide conflicting stories: some indicators show signs of improvements, while others suggest quality may have declined. Overall, however, there is little evidence of widespread declines in cognitive gains from schooling around the world. Furthermore, accounting for the potential decline in quality observed in some sources leaves the main results unchanged, because this decline appears to have been minor in comparison to the large observed increases in the quantity of schooling.

D.1. Trends in Education Quality: Comparison of Available Estimates

D.1.1. International Test Scores

The first set of available estimates on changes in quality come from international test scores, which have been increasingly conducted in most countries in the world since the 1990s-2000s. Drawing from various international sources, Angrist et al. (2021) compile test score results for 163 countries over the 2000-2017 period, 122 of which have at least one data point in the 2000s and another data point in the 2010s. The data suggest that education quality has remained broadly stable in most regions, despite some noticeable increases in quality observed in Sub-Saharan Africa and Latin America.

Figure D1 compares average test scores in the 2000s and 2010s for all countries with available data, based on the database of Angrist et al. (2021). Each data point corresponds to a test score in a given country, for a given education level (primary/secondary) and subject (mathematics/science/reading). All points are very close to the 45-degree line, suggesting that there has been little change in quality over the period. If anything, there has been a slight improvement in average quality: test scores have improved for 170 country-level-subject cells, while they have declined for 100.

For a more restricted number of countries, it is also possible to look at longer-run trends in education quality, based on the database of harmonized test scores compiled by Altinok, Angrist, and Patrinos (2018). Figure D2 plots the evolution of this indicator since 1970 for a selected number of high- and middle-income countries. The picture that arises is again one of remarkable stability, although some countries have undergone important long-run improvements in schooling quality, including Brazil, Chile, Iran, and South Korea.

D.1.2. Conditional Literacy

Test scores arguably provide the best available information, yet they suffer from a critical lack of historical depth for most countries in the world. To make a first step towards closing this gap, Le Nestour, Moscoviz, and Sandefur (2022) exploit information on literacy reported in the Demographic and Health Surveys and the Multiple Indicator Cluster Surveys. These surveys have repeatedly collected information on ability to read in many developing countries since 2000. The enormous advantage of these sources is that they cover adults, which allows tracking education quality across cohorts. This considerably expands the time period, given that the first cohorts covered by the data were born as early as the 1950s. To the best of my knowledge, this represents the only available approach to track historical trends in education quality in the developing world. Based on this, Le Nestour, Moscoviz, and Sandefur (2022) exploit repeated cross-sections to identify changes in education quality, defined as expected literacy at grade 5, across cohorts.

Figure D3 shows the main result of this exercise, comparing expected literacy at age 5 for cohorts born in 1950-1960s versus 1980-2000. The estimates of Le Nestour, Moscoviz, and Sandefur (2022) point to a clear decline in quality in a number of developing countries. In India, for instance, five years of schooling are found to be associated with about 50% of 1980-2000 cohorts being able to read, compared to 90% of 1950-1960 cohorts.

These results are insightful, but it is important to stress that they do not necessarily imply that the results presented in this paper should be revised downwards for at least four reasons.

First, ability to read is arguably a very partial and noisy measure of quality. For instance, Hermo et al. (2022) show that the decline of vocabulary knowledge in Sweden since the 1960s has been accompanied by a significant increase in logical reasoning skills, which can be rationalized by increasing labor market returns to the latter. In this context, relying solely on one dimension of quality (such as reading) could provide an inaccurate picture of changes in education quality.

Second, identifying trends in the quality of education from repeated cross sections of surveys requires explicitly modeling age, period, and cohort effects. This makes the results much more sensitive to methodological choices, measurement error, and potential sampling differences across survey waves, all of which can be particularly acute in developing countries. The results presented on South Africa in the next section suggest that this is an important concern.

Third, such estimates are not immune to standard problems associated with causal identification (which is also true of test scores). An important source of bias is that improvements in access to schooling in the developing world have been overwhelmingly concentrated among children coming from low-income and lower-educated families (Gethin, 2023). As a result, lower performance among newly educated cohorts may primarily be the result of greater cognitive and socioeconomic barriers to learning, rather than to changes in the value added of schooling.

Finally, changes in average performance may not necessarily imply lower returns to schooling. Even if newly educated cohorts may have lower levels of cognitive skills, economic returns to schooling for them may still be equal, or even greater (as suggested by the IV returns to schooling presented in the main text), than returns for the rest of the population. What matters is not whether newly skilled workers have lower or higher levels of skills, but instead what are the returns to increased access to schooling for this specific subset of the population. Put differently, differences in average skills may be very different from differences in marginal returns to skill.

D.1.3. Country-Specific Sources: Insights from South Africa

Despite these limitations, the cohort-based approach developed by Le Nestour, Moscoviz, and Sandefur (2022) has another advantage: it can be extended to other countries and data sources reporting information on education quality. Indeed, the DHS/MICS are not the only surveys recording information on adult skills. Applying the same methodology to other datasets and country-specific sources provides a fruitful avenue for future research.

While engaging in such a vast data collection and harmonization effort goes beyond the objective of this paper, I draw on previous work (Gethin, 2022) to document long-run trends in quality in one context: South Africa. Indeed, the General Household Survey has collected detailed information since 2009 on adults' ability to perform six basic operations: writing one's name, reading newspapers and other documents, filling in a form, writing a letter, calculating how much change should be received when buying something, and reading road signs. Information is collected for each household member, with four values ranging from "No difficulty" to "Unable to do."

Drawing on these cross-sections, I run simple regressions relating scores on these indicators to completed years of schooling, controlling for gender, race, province of residence, and survey year fixed effects. Regressions are run by decade of birth to capture cohort changes in education quality. I normalize each dependent variable to range from 0 to 1. Coefficients of interest can then be interpreted as expected literacy obtained from an additional year of schooling.

Figure D4 plots the resulting evolution of coefficients by decade of birth. Despite some fluctuations and differences in expected gains across items, education quality is estimated to have remained extremely stable from the 1940s to the 1980s. On average, a year of schooling is associated with a 10 percentage point increase in literacy.

This result is puzzling, given that South Africa is one of the countries with the largest estimated decline in quality in the Le Nestour, Moscoviz, and Sandefur (2022) data. Indeed, conditional literacy at grade 5 is found by the authors to have decreased by as much as 20 percentage points, from about 70% to 50% (see figure D3).

This conflicting evidence suggests that much more research is needed before reaching decisive conclusions on trends in education quality in the developing world. Test scores are perhaps the best data available, but they do not exist before the 2000s in many countries. Cohort trends in literacy are arguably a promising indicator, but data sources and methodologies remain to be further tested and compared.

D.1.4. Returns to Schooling Among U.S. Migrants

A last piece of evidence comes from returns to schooling among U.S. migrants. Schoellman (2012) argues that differences in returns to schooling among U.S. migrants originating from different countries provides a good proxy for education quality, because it captures income gains from schooling for individuals having been educated in different countries but working in the same labor market. For instance, returns to schooling are expected to be higher among Swedish migrants than among Congolese migrants, as differences in educational attainment reflect greater differences in accumulated human capital in the former group than in the latter. Schoellman (2012) provides evidence that this indicator is a good proxy for education quality, strongly correlating with GDP per capita and available test scores (see also Rossi, 2022).

The advantage of returns to schooling among migrants is that they can be estimated for an even greater number of countries than cohort trends in literacy studied in Le Nestour, Moscoviz, and Sandefur (2022). Pooling several waves of U.S. censuses, it is also possible to estimate returns to schooling for different cohorts of migrants. Although this analysis is evidently not devoid of limitations—in particular small sample sizes and potential differential selection into schooling across cohorts of a given country—, it can still hopefully shed light on broad long-run trends.

I pool 1980, 1990, and 2000 U.S. censuses, together with all American Community Surveys from 2001 to 2021. I restrict the sample to individuals aged 25 to 65 with positive earned income, who were born outside of the U.S. between 1950 and 1980, and arrived in the U.S. after age 20. I then run the following regressions:

$$y_{icyt} = \zeta_{cy} s_{icyt} + X_{icyt}\beta_{cy} + \mu_t \tag{51}$$

With y_{icyt} the log of total yearly earned income of individual i born in country c in decade y (1950s, 1960s, 1970s, or 1980s) and observed in year t. s_{icyt} is completed years of schooling, X_{icyt} are control variables (gender, state of residence, and year of immigration), and μ_t are census/ACS year fixed effects. The parameter of interest is ζ_{cy}, the return to a year of schooling for individuals born in country c in decade y. If education quality has declined substantially, then we should expect ζ_{cy} to have declined over time: a year of schooling should deliver greater returns for migrants born in the 1950s than for migrants born in the 1980s. I run this regression separately for each country of origin × decade of birth cell.

The results of this exercise are presented in figure D5, which plots population-weighted averages of the estimated returns to schooling by world region of birth and decade of birth. Returns to schooling are lowest among migrants from Latin America and Sub-Saharan Africa and highest among migrants from Europe and the Anglosphere (Canada, Australia, New Zealand, United Kingdom). There are fluctuations across decades, but no clear trend in quality in most regions. Returns have fluctuated at about 4-6% per year of schooling among Latin American migrants, compared to 9-11% among European and Anglosphere natives. The world average varies from 7% to 9% with no clear long-run evolution.[32] This suggests again that changes in education quality are unlikely to play a substantial role in affecting the results presented in this paper.

D.2. A Quantification Exercise

While it remains unclear which data source should be preferred, it is still useful to test how sensitive are my main findings to accounting for the potential decline in quality documented in Le Nestour, Moscoviz, and Sandefur (2022). This is somewhat of a heroic task, because it requires (1) extrapolating cohort trends to cover education quality for the entire 1980-2019 working-age populations (2) putting a monetary value on literacy, to build measures of quality-adjusted years of schooling, and (3) extrapolating changes in quality to countries with no

[32] It is also interesting to investigate differences between cohort trends in returns to schooling among migrants and in literacy rates estimated by Le Nestour, Moscoviz, and Sandefur (2022). The raw cross-country correlation between changes in returns and changes in literacy from the 1960s to the 1980s cohorts is 0.33. This suggests that both sources tell a broadly similar story on which countries have seen education quality decline or improve most.

available data. This section represents an exploratory attempt at doing so.

D.2.1. Methodological Framework

Constructing estimates of quality-adjusted years of schooling requires mapping education quality into equivalent years of schooling. Following the existing literature, I consider the following standard extension of the Mincer-type human capital stock (e.g., Hanushek, Ruhose, and Woessman, 2017):

$$h = \exp(r_L L + r_Q Q) \tag{52}$$

With r_L the return to a year of schooling, L average years of schooling, r_Q the return to education quality, and Q an indicator of education quality. The objective is to convert a change in quality from Q to $\tilde{Q}$ into an equivalent change in years of schooling from L to $\tilde{L}$. This equivalence satisfies:

$$\exp(r_L L + r_Q \tilde{Q}) = \exp(r_L \tilde{L} + r_Q Q) \tag{53}$$

Rearranging:

$$\tilde{L} = L - \frac{r_Q}{r_L}(Q - \tilde{Q}) \tag{54}$$

Calculating quality-adjusted changes in years of schooling thus requires data on changes in education quality $(Q - \tilde{Q})$, as well as the relative value of schooling quality (r_Q) compared to schooling quantity (r_L). I now turn to estimating each of these two components.

D.2.2. Estimation of Global Trends in Conditional Literacy

The first step is to estimate $(Q - \tilde{Q})$, the evolution of quality of schooling for the working-age population from 1980 to 2019. The database of Le Nestour, Moscoviz, and Sandefur (2022) provides information on literacy at grade 5 in 86 countries for two cohorts born during the 1952-1999 period (see Le Nestour, Moscoviz, and Sandefur, 2022, Table 7). Starting from these two data points by country, I estimate average conditional literacy for the working-age population.

First, I divide all figures by 5, so that the indicator corresponds to expected literacy per year of education. This ensures that the education quality indicator is comparable to years of schooling.

Second, I linearly interpolate and extrapolate this indicator backwards and forwards, to cover all cohorts born from 1915 to 1994. This is a very conservative assumption: it amounts to considering that education quality continued to decline at the same pace after the last cohort observed, and was already declining at the same pace from 1915 until the first cohort observed. This is unlikely to be true, given evidence documented above on the stability or even rise of education quality in many countries since the 2000s.

Third, I construct measures of average education quality of the working-age population. To do this, I average the indicator over all cohorts aged 25 to 65 in a given year, weighted by the population of each cohort. Data on population by age is taken from the United Nations' World Population Prospects. The result is an indicator of education quality covering the working-age population of each country from 1980 to 2019, corresponding to average expected literacy per year of schooling.

Finally, in the absence of data for the rest of the world, I impute the indicator for missing countries using three polar scenarios. The benchmark scenario assumes that education quality in missing countries has declined at the same pace as the average decline observed over the 86 countries. The upper bound assumes that it has not declined. The lower bound assumes that it has declined at the speed of India, that is, at a very fast pace (see figure D3). I view this last case as an extreme and implausible scenario, given above-mentioned evidence on the stability or rise of test scores in many countries.

Figure D6 compares education quality of the working-age population in 1980 and 2019 for countries with available data. The overall pattern and ranking of countries is similar to the one visible in figure D3. However, the change in quality appears less dramatic, because this figure compares education quality for the overall population rather than across cohorts. In India, for instance, literacy per year of schooling declined by about 8 percentage points, from 22 to 14.

D.2.3. Estimation of Returns to Literacy

The second step is to estimate r_Q/r_L, the returns to literacy relative to a year of schooling. This requires data on personal income, years of schooling, and literacy at the individual level. I was able to find four high-quality surveys covering these three variables: the Brazilian 2015 PNAD survey, the Indonesian 1998 SUSENAS survey, the Pakistani 2018 HIES survey, and the South African 2019 GHS survey. In each of these four countries, I estimate the relative returns to literacy by running two regressions: a regression relating the log of total personal income to literacy, and a regression relating the log of total personal income to years of schooling, controlling for gender, potential experience, and potential experience squared in each case. I restrict the sample to workers with either no schooling or basic education, to make sure that the two estimates are comparable (nearly all workers with more than basic education are literate).

The results are presented in table D1. Returns to schooling range from 3% to 8% per year of basic education, while returns to literacy range from 18 to 39 log points. The ratio between the two coefficients, corresponding to r_Q/r_L, is very similar across countries, ranging from 5 in Pakistan to about 6.5 in Indonesia. I take a value of 6 to construct measures of quality-adjusted years of schooling in what follows. This amounts to assuming that moving the entire population from being illiterate to literate is equivalent to increasing average schooling by 6 years.

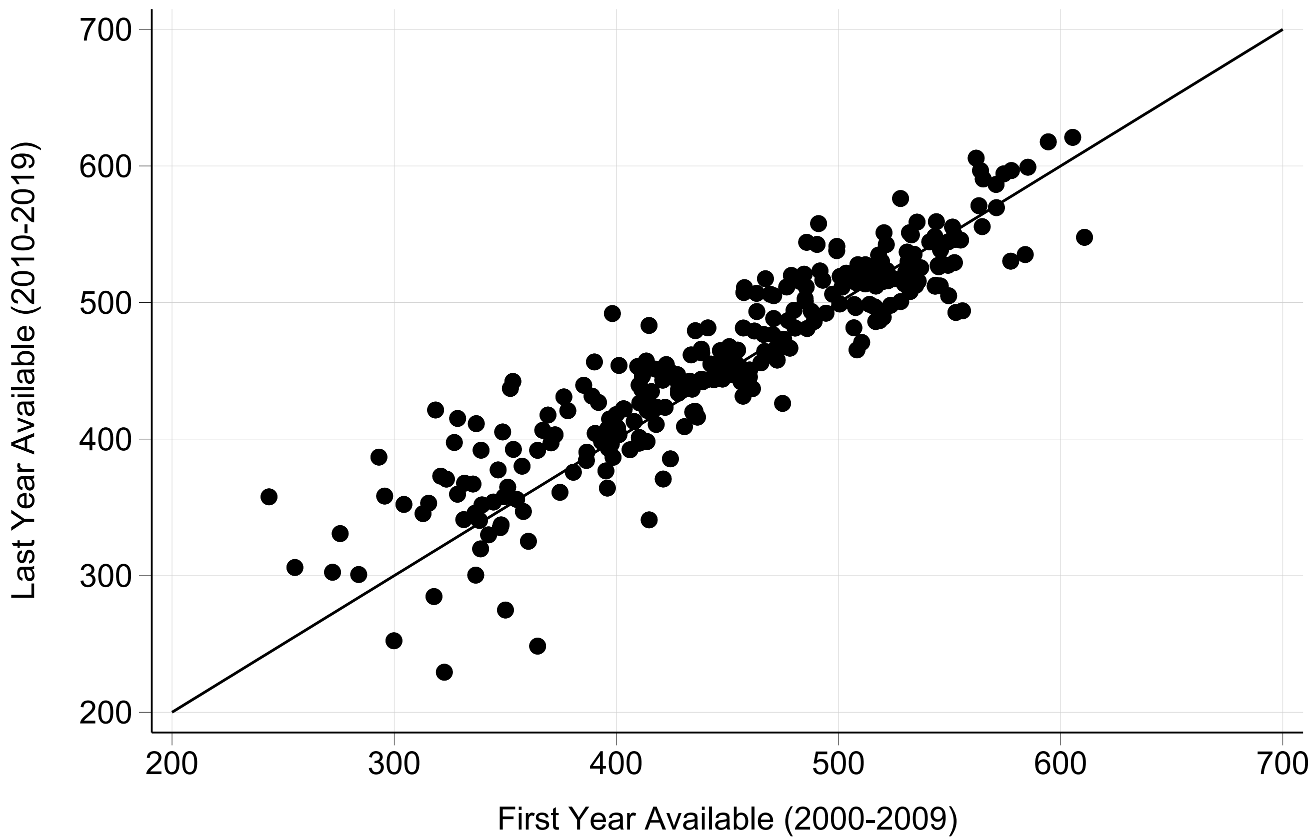
D.2.4. Results

Having estimated changes in education quality and its price relative to education quantity, one can now construct measures of quality-adjusted years of schooling. In practice, I set 1980 as the benchmark year, and adjust estimates of average years of schooling in all other years from 1981 to 2019 so that they reflect the quality observed in 1980. For instance, quality-adjusted years of schooling in 2019 are calculated as $\tilde{L}_{2019} = L_{2019} - \frac{r_Q}{r_L}(Q_{2019} - Q_{1980})$, with L_{2019} unadjusted years of schooling observed in 2019, $\frac{r_Q}{r_L} = 6$, and $Q_{2019} - Q_{1980}$ the change in expected literacy per year of schooling from 1980 to 2019. This approach thus amounts to "deflating" years of schooling observed from 1981 to 2019 to express them in 1980 equivalents.

Figure D7 compares the evolution of average years of schooling in the world as a whole, before and after adjusting for changes in education quality. The unadjusted indicator rose from 5 to 8.5. Years of schooling expressed in 1980 equivalents rose from 5 to 8.1-8.3. Adjusting for education quality thus reduces average years of schooling today by at most 0.5 years (or 6%), and the overall increase in education since 1980 by at most 0.4 years (or 9%).

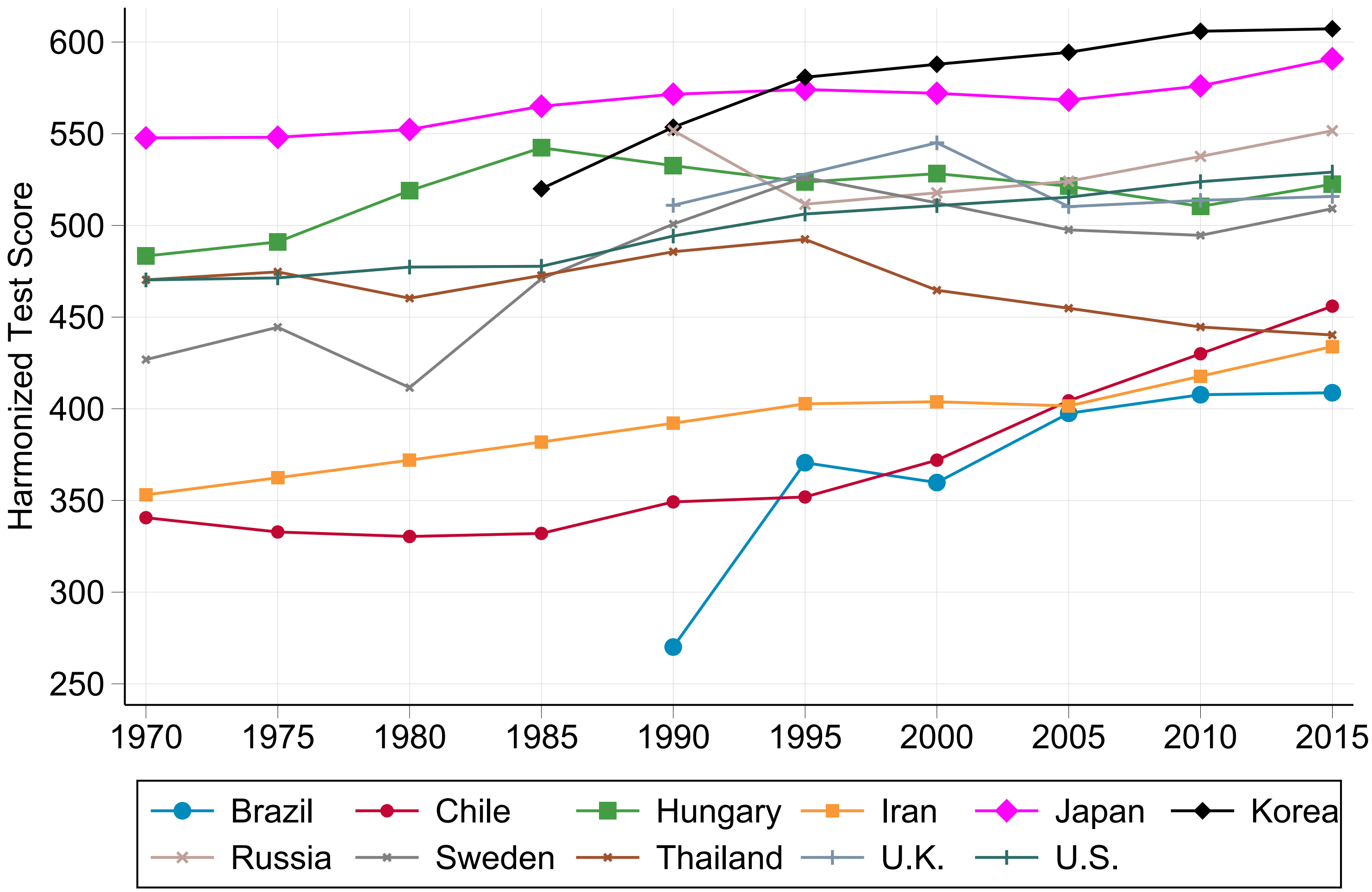
Figure D8 compares the share of growth explained by global income percentile before and after making the lower bound adjustment (assuming that education quality declined as fast as in India for all countries with missing data). The two lines are barely distinguishable: even under strong assumptions on the decline in education quality, the main result remains almost unchanged. Overall, the share of growth explained by education declines by about 2 to 6 percentage points depending on the percentile considered, with the greatest changes observed at the upper-middle of the income distribution. The results presented in this paper thus appear to be strongly robust to potential changes in education quality observed since 1980.

Figure D1 – Harmonized Test Scores by Country: 2000-2009 vs. 2010-2019



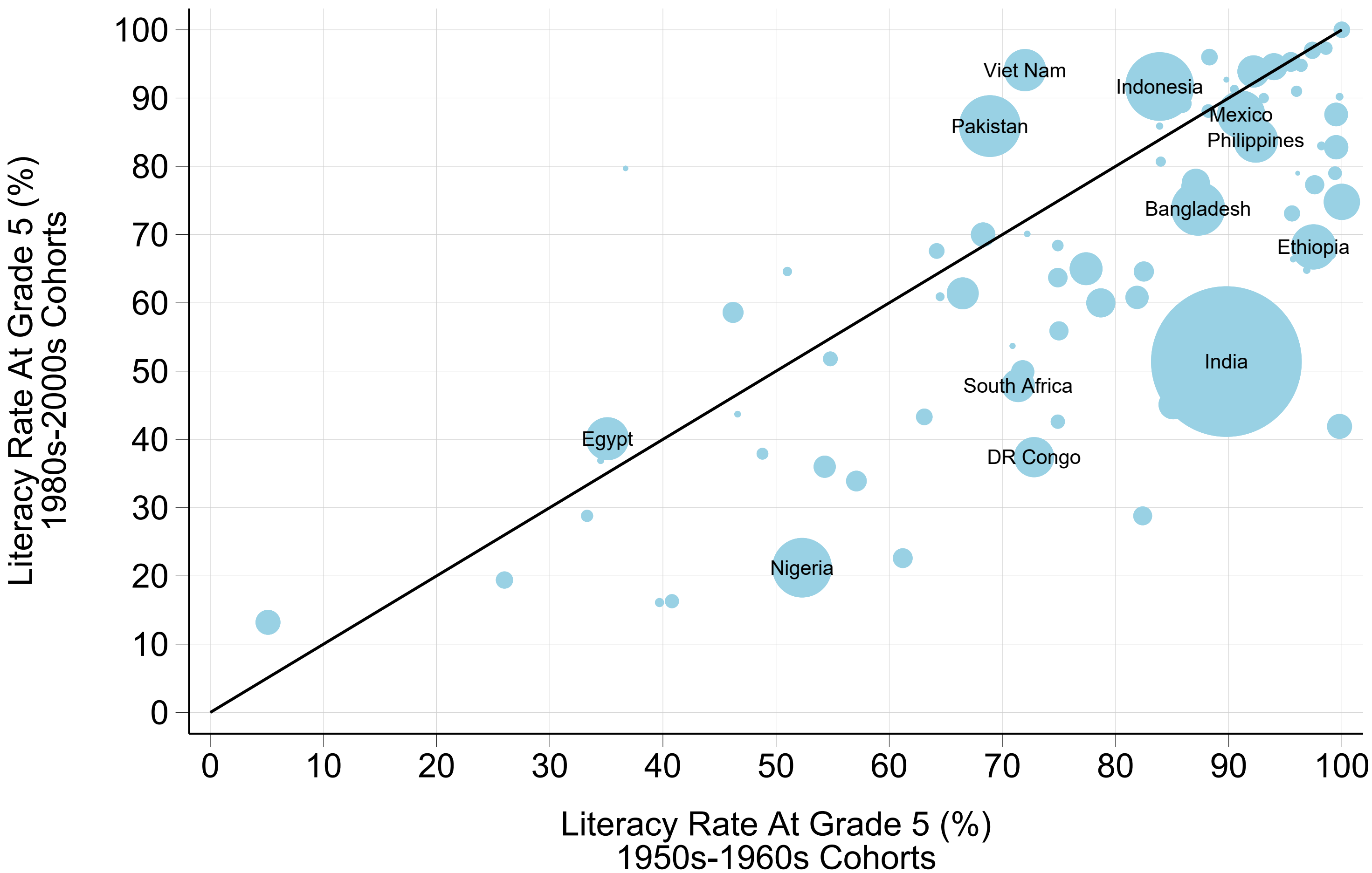
Source: Angrist et al. (2021). Each point corresponds to a test score reported for a given country × education level (primary/secondary) × subject (maths/science/reading).

Figure D2 – Long-Run Trends in Test Scores in Selected Countries, 1970-2015



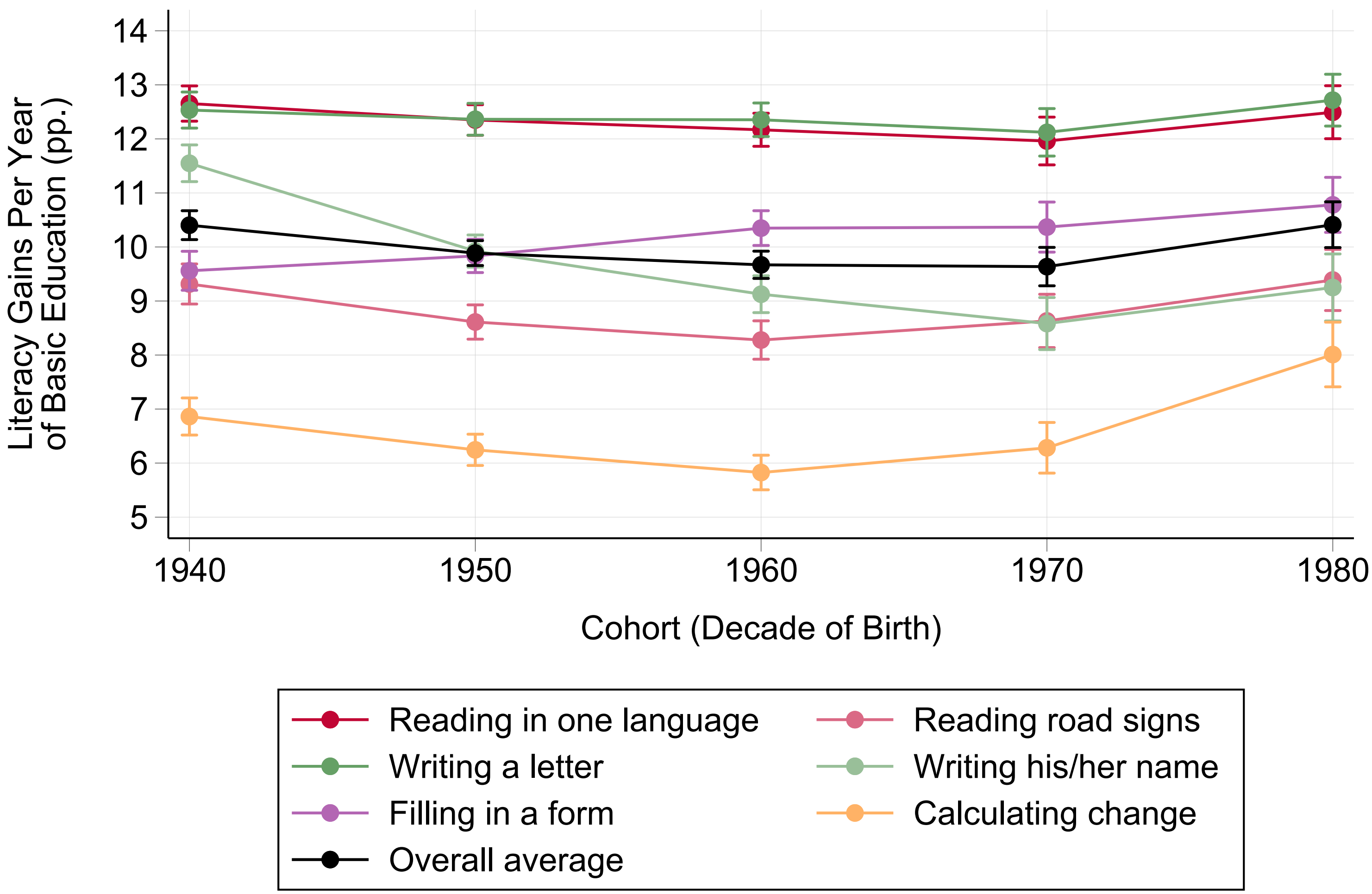
Source: Altinok, Angrist, and Patrinos (2018).

Figure D3 – Literacy at Grade 5: 1950-1960 versus 1980-2000 Cohorts



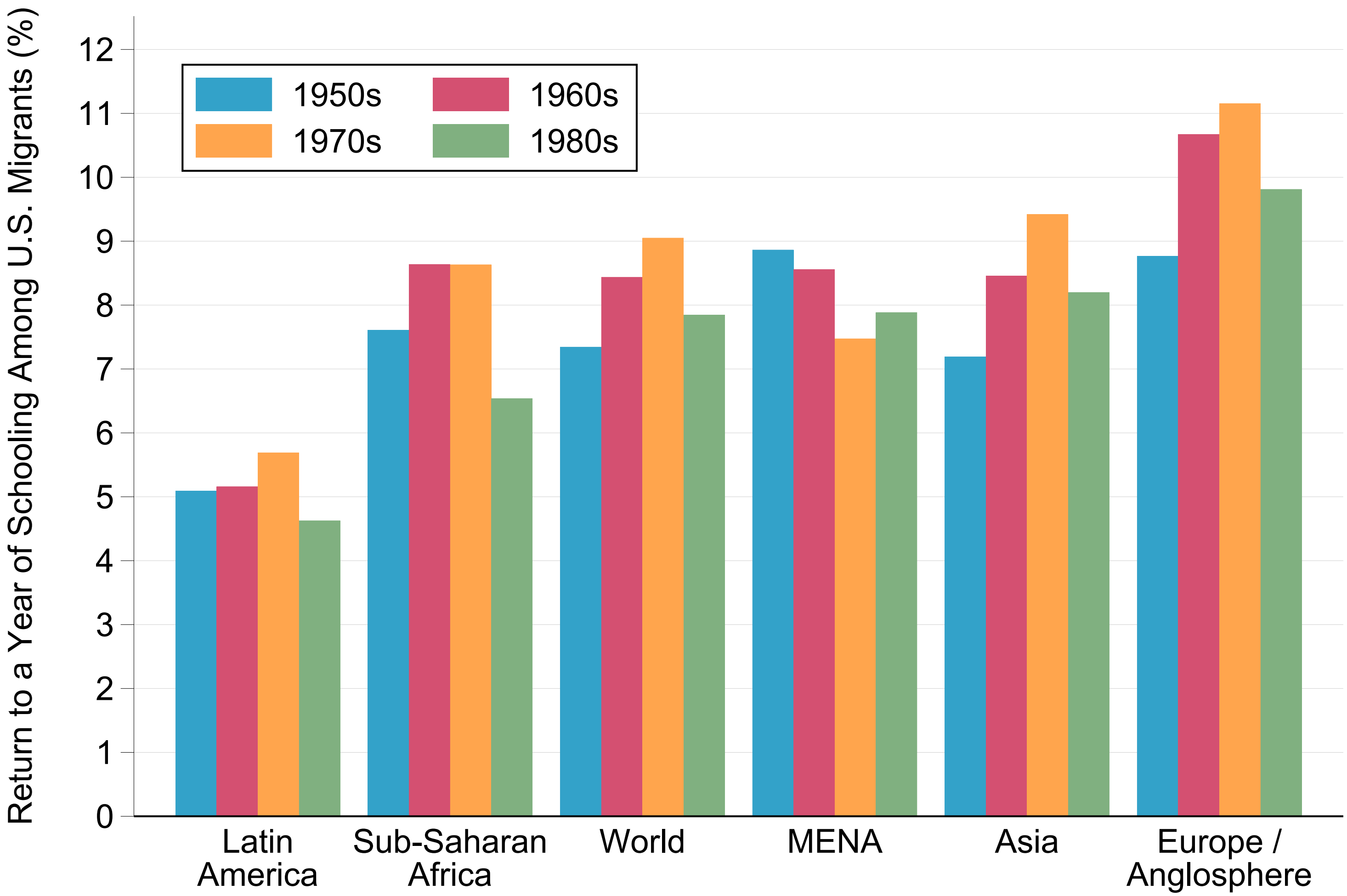
Source: Le Nestour, Moscoviz, and Sandefur (2022).

Figure D4 – Long-Run Trends in Education Quality in South Africa:
Cognitive Gains Per Year of Basic Education, 1940-1980 Cohorts



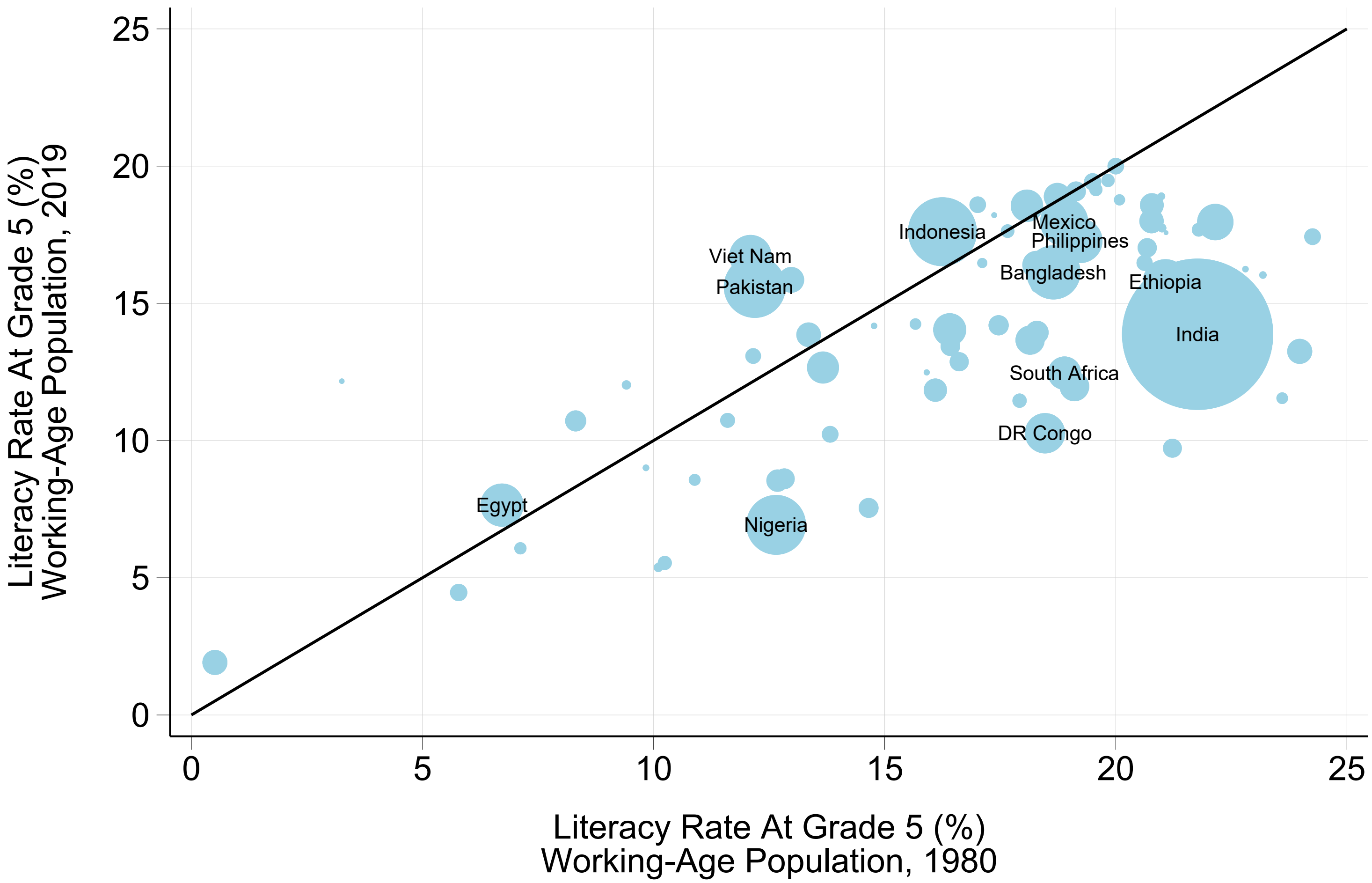
Note: author's calculations using South African General Household Surveys (2009-2013). Each data points corresponds to the coefficient of a regression of corresponding literacy scores (0-1) on years of schooling, restricting the sample to individuals with 0 to 6 years of education.

Figure D5 – Trends in Returns to Schooling Across Cohorts of U.S. Migrants



Note: author's calculations using U.S. censuses and American Community Surveys. Each bar corresponds to the population-weighted average of returns to schooling estimated for a given country of origin × decade of birth cell.

Figure D6 – Literacy Gains Per Year of Schooling: 1980 versus 2019 Working-Age Population



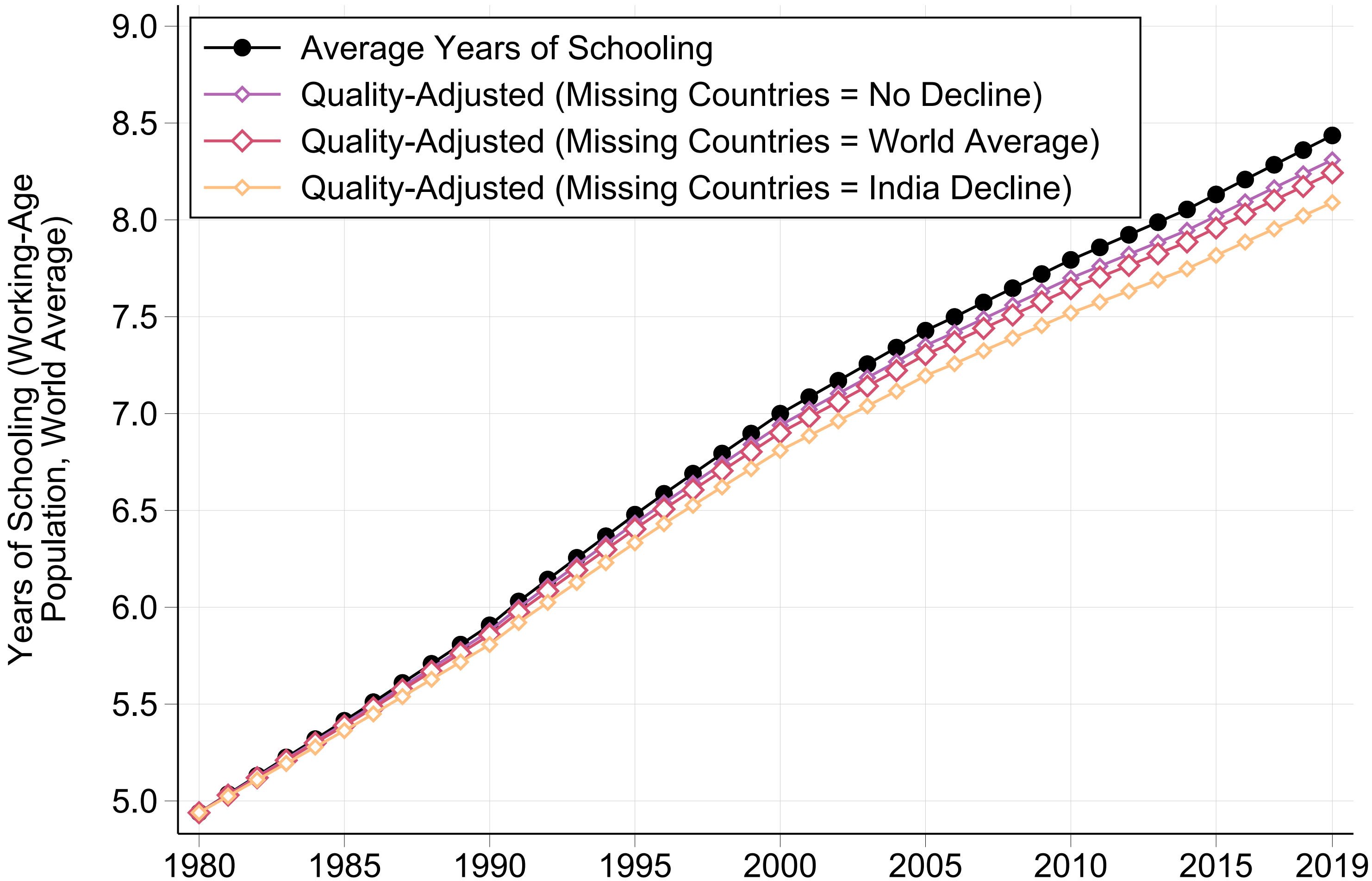
Note: author's estimation combining data from Le Nestour, Moscoviz, and Sandefur (2022) and United Nations World Population Prospects.

Table D1 – Returns to Literacy

	Brazil	Indonesia	Pakistan	South Africa
Return to Literacy	0.39***	0.31***	0.31***	0.18***
	(0.01)	(0.01)	(0.01)	(0.05)
Return to Schooling	0.08***	0.05***	0.06***	0.03***
	(0.00)	(0.00)	(0.00)	(0.01)
Literacy / Schooling	5.10	6.53	5.03	5.66

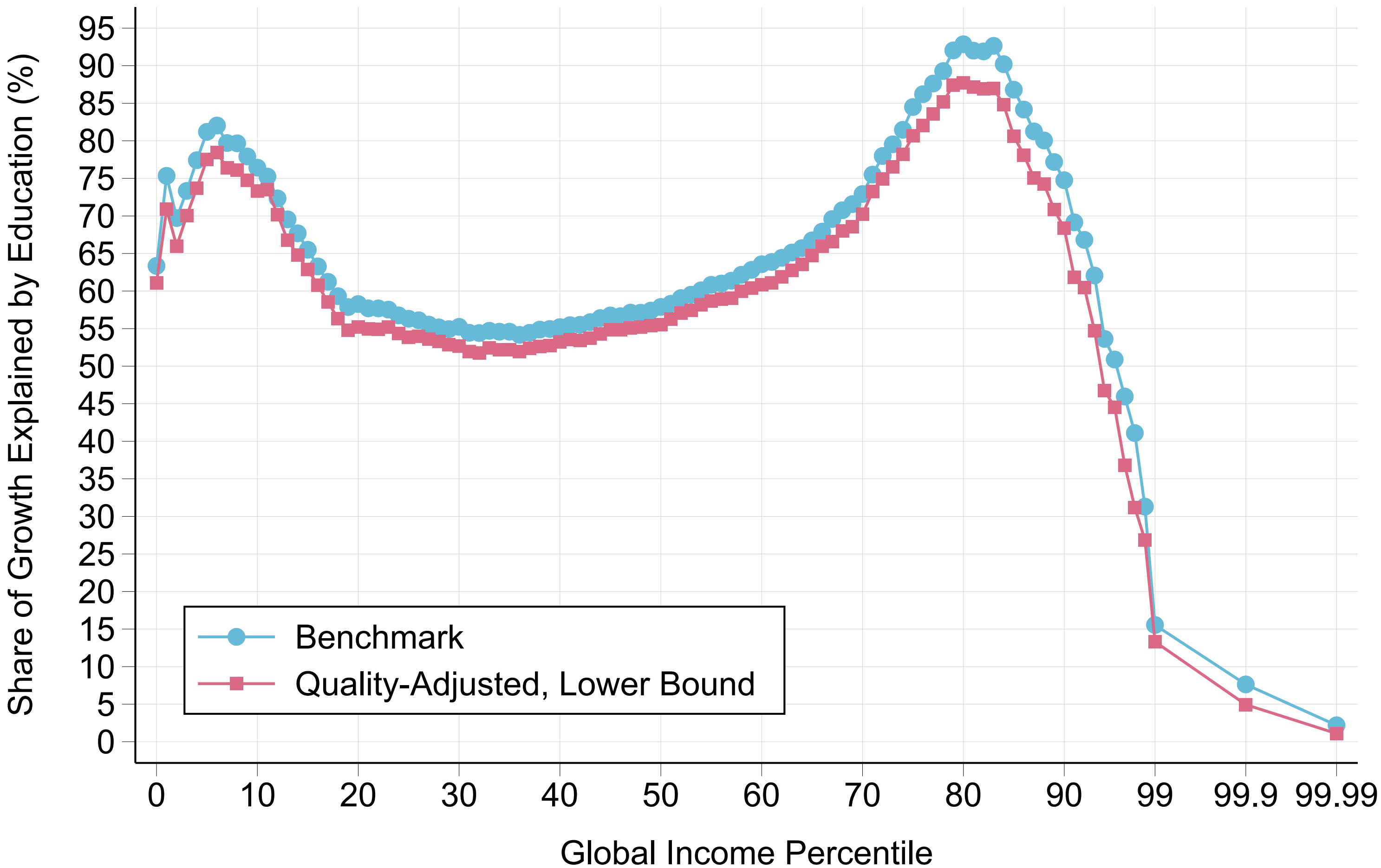
Notes. The table reports estimates of returns to literacy, returns to schooling, and the ratio between the two. The coefficient on literacy corresponds to a regression of the log of personal income on literacy; the coefficient on years of schooling corresponds to a separate regression of the log of personal income on years of schooling. Both regressions control for gender, potential experience, and potential experience squared in each country. Data sources: 2015 Brazil PNAD survey, 1998 Indonesia SUSENAS survey, 2018 Pakistan HIES survey, 2019 South Africa GHS survey. Standard errors in parenthesis. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Figure D7 – Global Average Years of Schooling: Unadjusted versus Quality-Adjusted Estimates Using Cohort Conditional Literacy Rates



Note: unadjusted estimates correspond to average years of schooling of the world working-age population. Quality-adjusted estimates correct years of schooling for the decline in education quality estimated by Le Nestour, Moscoviz, and Sandefur (2022), so that years of schooling are expressed in 1980 equivalents throughout the period. Upper bound: education quality assumed to have remained constant for all countries with no data from Le Nestour, Moscoviz, and Sandefur (2022). Benchmark: countries with missing data are attributed the average yearly change in quality estimated across all countries covered. Lower bound: countries with missing data are attributed the decline in quality observed in India.

Figure D8 – Share of Growth Explained by Education: Benchmark Versus Lower Bound on Decline in Education Quality



Note: Quality-adjusted estimates correct years of schooling for the decline in education quality estimated by Le Nestour, Moscoviz, and Sandefur (2022), so that years of schooling are expressed in 1980 equivalents throughout the period. Countries with missing data are attributed the decline in quality observed in India.

E. Data Appendix: Survey Microdata

The survey microdata covering education and earnings in 150 countries used in this paper come from four main data sources.

ILO Microdata The main data source is a set of harmonized household surveys that were collected and compiled by the International Labor Organization. The ILO database covers over 1,400 surveys fielded in 136 countries from 1990 to 2022. In the main analysis of this paper, I use the last survey available in each country. However, I also exploit historical surveys in the analysis of backward versus forward accounting presented in section 6.2. The database presents itself as a single harmonized microfile. The main variables are country, year, household ID, sample weight, wage income (from main job, second job, and all jobs combined), self-employment income (from main job, second job, and all jobs combined), age, gender, education, labor force participation, occupation (ISCCO-08), industry, and rural-urban location. I define personal income as the sum of all wage and self-employment income received by an individual. I drop all zeros and missing values, so that the sample is restricted to all individuals with strictly positive personal income.

European Statistics on Income and Living Conditions Although the ILO microdata do cover European countries, the coding of educational attainment is broader than in the original microfiles, so I decide to rely on my own data collection. The European Statistics on Income and Living Conditions (EU-SILC) cover detailed information on personal income and education in 32 countries every year from 2003 to 2020. I harmonize EU-SILC surveys in the same way as those of the ILO, defining personal income as the sum of individual wage and mixed income. I then replace all ILO surveys by this microfile, with the exception of France, Portugal, and Switzerland, for which the ILO provides national labor force surveys of even better quality.

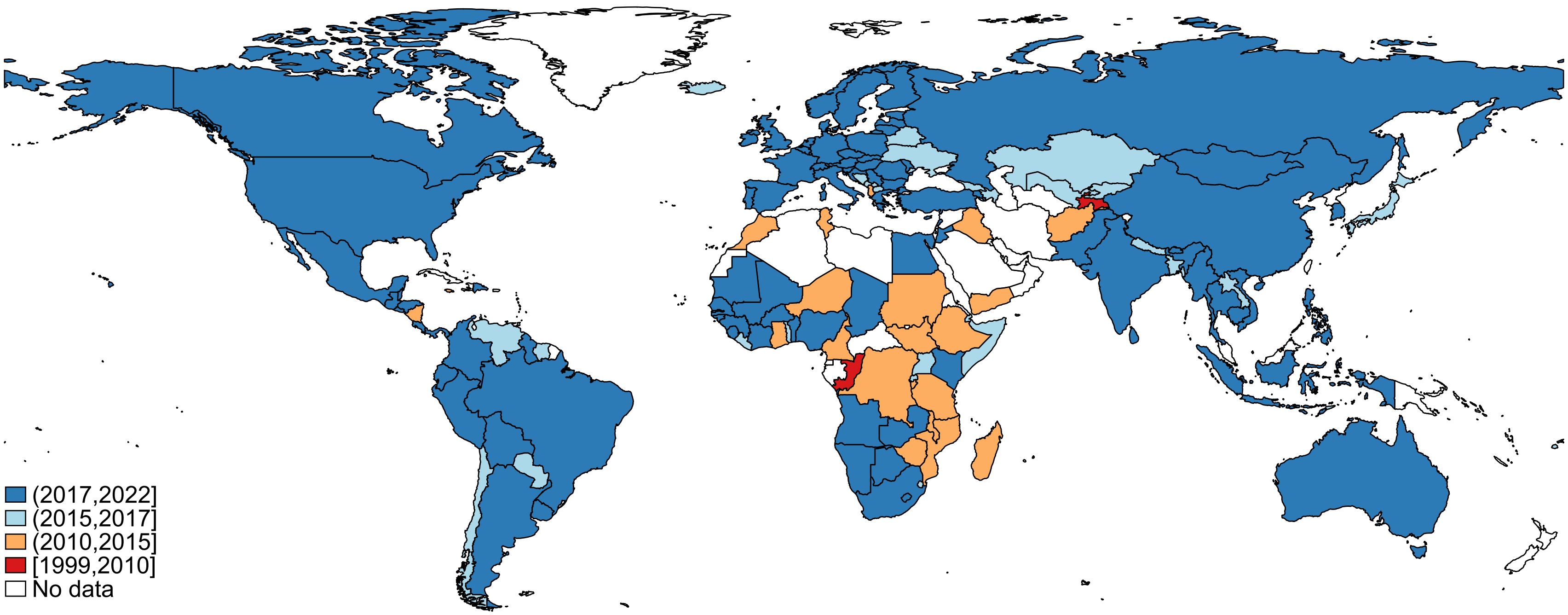
Life in Transition Survey For 10 Eastern European and Central Asian countries not covered by the ILO, I rely on the Life in Transition Survey (LITS). These are Azerbaijan, Belarus, Georgia, Kyrgyzstan, Kosovo, Kazakhstan, North Macedonia, Montenegro, Ukraine, and Uzbekistan, for which labor force or household living standards surveys are unfortunately not publicly accessible at the time of writing. The LITS is far from being ideal, with sample sizes of only 3,000-5,000 in each country, yet it is to the best of my knowledge the only data source available to measure individual incomes and education. I use the last wave of the LITS, fielded in 2016, which I harmonize in the same way as the ILO.

Country-Specific Surveys Finally, I collect and harmonize surveys from country-specific data portals to cover 13 additional countries: China, Iraq, India, Japan, Mozambique, Morocco, Russia, Somalia, South Africa, South Korea, South Sudan, Tunisia, and the United States.

For seven countries, I was available to find and harmonize a high-quality survey providing detailed information on individual incomes and education. This type of survey was available for China (2018 Chinese Household Income Project), India (2019 Periodic Labor Force Survey), Russia (2019 Russia Longitudinal Monitoring Survey), South Korea (2019 Korean Labor and Income Panel Study), Tunisia (2014 Labor Force Survey), South Africa (2019 General Household Survey), and the United States (2019 Current Population Survey).

For the remaining six countries, I rely on surveys of lower quality or only providing information on household expenditure. For Japan, in the absence of better publicly available data, I use the 2017 general household survey, which does cover individual income and education but has a small sample size (about 1,000). I use household income and expenditure surveys for Iraq (Household Socio-Economic Survey), Mozambique (Inquérito aos orcamentos familiares), Morocco (Household Expenditure Survey), Somalia (High Frequency Survey), and South Sudan (High Frequency Survey), which provide information on individual employment and education, as well as total household expenditure, but not on individual incomes. In the absence of better information, I proxy personal income by splitting equally household expenditure among adults in employment, excluding unemployed or inactive individuals as well as children.

Figure E1 – Survey Data Coverage: Year Covered by Each Survey



Notes. Colored countries are those covered by the survey microdata. Colors correspond to the year during which each survey was fielded.

Table E1 – Survey Data Sources

Country	Source	Survey Year
Europe		
Albania	Living Standards Survey	2012
Austria	EU Statistics on Income and Living Conditions	2019
Belarus	Life in Transition Survey	2016
Belgium	EU Statistics on Income and Living Conditions	2019
Bosnia and Herzegovina	Labour Force Survey	2016
Bulgaria	EU Statistics on Income and Living Conditions	2019
Croatia	EU Statistics on Income and Living Conditions	2019
Czechia	EU Statistics on Income and Living Conditions	2019
Denmark	EU Statistics on Income and Living Conditions	2019
Estonia	EU Statistics on Income and Living Conditions	2019
Finland	EU Statistics on Income and Living Conditions	2019
France	Employment Survey	2019
Germany	EU Statistics on Income and Living Conditions	2019
Greece	EU Statistics on Income and Living Conditions	2019
Hungary	EU Statistics on Income and Living Conditions	2019
Iceland	EU Statistics on Income and Living Conditions	2018
Ireland	EU Statistics on Income and Living Conditions	2019
Italy	Labour Force Survey	2019
Latvia	EU Statistics on Income and Living Conditions	2019
Lithuania	EU Statistics on Income and Living Conditions	2019
Luxembourg	EU Statistics on Income and Living Conditions	2019

Malta	EU Statistics on Income and Living Conditions	2019
Moldova	Labour Force Survey	2019
Montenegro	Life in Transition Survey	2016
Netherlands	EU Statistics on Income and Living Conditions	2019
North Macedonia	Life in Transition Survey	2016
Norway	EU Statistics on Income and Living Conditions	2019
Poland	EU Statistics on Income and Living Conditions	2019
Portugal	Employment Survey	2019
Romania	EU Statistics on Income and Living Conditions	2019
Russia	Russia Longitudinal Monitoring Survey	2019
Serbia	Labour Force Survey	2019
Slovakia	EU Statistics on Income and Living Conditions	2019
Slovenia	EU Statistics on Income and Living Conditions	2019
Spain	EU Statistics on Income and Living Conditions	2019
Sweden	EU Statistics on Income and Living Conditions	2019
Switzerland	Labour Force Survey	2019
Ukraine	Life in Transition Survey	2016
United Kingdom	Labour Force Survey	2018
Northern America		
Canada	Labour Force Survey	2019
USA	Current Population Survey	2019
Latin America		
Argentina	Permanent Household Survey, Urban	2019
Barbados	Survey on Living Conditions	2016
Belize	Labour Force Survey	2019

Bolivia	Continuous Employment Survey	2019
Brazil	Continuous National Household Sample Survey	2019
Chile	National Survey on Socio-Economic Conditions	2017
Colombia	Integrated Household Survey	2019
Costa Rica	National Household Survey	2019
Dominican Republic	Continuous National Labour Force Survey	2019
Ecuador	National Survey on Employment	2019
El Salvador	Multi-purpose Household Survey	2019
Guatemala	Monthly Employment and Income Survey	2019
Guyana	Labour Force Survey	2019
Honduras	Continous Multi-Purpose Household Survey	2019
Jamaica	Labour Force Survey	2014
Mexico	National Occupation and Employment Survey	2019
Nicaragua	National Household Survey on Measuring Living Conditions	2014
Panama	Labour Market Survey	2019
Paraguay	Continous Household Survey	2017
Peru	National Household Survey	2019
Suriname	Survey on Living Conditions	2016
Trinidad and Tobago	Continuous Sample Survey of the Population	2016
Uruguay	Continous Household Survey	2019
Venezuela	Household Sample Survey	2017
Asia		
Afghanistan	Households Living Conditions Survey	2014
Australia	Household, Income and Labour Dynamics Survey	2019
Bangladesh	Labour Force Survey	2017

Bhutan	Labour Force Survey	2019
Brunei Darussalam	Labour Force Survey	2014
Cambodia	Labour Force Survey	2019
China	China Household Income Project	2018
Fiji	Employment, Unemployment Survey	2016
India	Periodic Labour Force Survey	2019
Indonesia	National Labour Force Survey	2019
Japan	General Social Survey	2017
Kazakhstan	Life in Transition Survey	2016
Kosovo	Life in Transition Survey	2016
Kyrgyzstan	Life in Transition Survey	2016
Lao	Labour Force Survey	2017
Maldives	Household Income and Expenditure Survey	2019
Mongolia	Labour Force Survey	2019
Myanmar	Labour Force Survey	2019
Nepal	Labour Force Survey	2017
Pakistan	Labour Force Survey	2019
Philippines	Labour Force Survey	2018
South Korea	Korean Labor and Income Panel Study	2019
Sri Lanka	Labour Force Survey	2018
Tajikistan	Living Standards Survey	2009
Thailand	Household Socio-Economic Survey	2019
Timor-Leste	Labour Force Survey	2016
Tonga	Labour Force Survey	2018
Uzbekistan	Life in Transition Survey	2016

Vietnam	Labour Force Survey	2019
Middle East and North Africa		
Armenia	Household Labour Force Survey	2019
Azerbaijan	Life in Transition Survey	2016
Cyprus	EU Statistics on Income and Living Conditions	2019
Egypt	Labour Force Sample Survey	2018
Georgia	Life in Transition Survey	2016
Iraq	Household Socio-Economic Survey	2012
Jordan	Employment and Unemployment Survey	2019
Lebanon	Labour Force Survey	2019
Morocco	Household Expenditure Survey	2014
Palestine	Labour Force Survey	2019
Sudan	Household Survey	2011
Tunisia	Labor Force Survey	2014
Turkey	Household Labour Force Survey	2019
Yemen	Labour Force Survey	2014
Sub-Saharan Africa		
Angola	Employment Survey	2019
Benin	Integrated Survey of Household Living Conditions	2018
Botswana	Multi-Topic Household Survey	2019
Burkina Faso	Regional Integrated Survey on Employment and the Informal Sector	2018
Burundi	Living Standards Survey	2014
Cabo Verde	Continuous Multi-Objective Survey	2015
Cameroon	Household Survey	2014
Chad	Modular and Integrated Household Survey on Living Conditions	2018

Comoros	National Survey on Employment and the Informal Sector	2014
Côte d’Ivoire	National Survey on the Employment Situation	2019
Democratic Republic of the Congo	Survey on Employment and household’s living conditions	2012
Djibouti	Djiboutian Household Survey	2017
Eswatini	Labour Force Survey	2016
Ethiopia	National Labor Force Survey	2013
Gambia	Labour Force Survey	2018
Ghana	Labour Force Survey	2015
Guinea	National Survey on Employment and the Informal Sector	2019
Guinea-Bissau	Harmonized Survey on Household Living Conditions	2018
Kenya	Household Budget Survey	2019
Lesotho	Labour Force Survey	2019
Liberia	Labour Force Survey	2017
Madagascar	National Survey on Employment and the Informal Sector	2015
Malawi	Labour Force Survey	2013
Mali	Continous Household Employment Survey	2018
Mauritania	Living Standards Survey	2019
Mauritius	Continuous Multi-Purpose Household Survey	2019
Mozambique	Inquérito aos orçamentos familiares	2014
Namibia	Labour Force Survey	2018
Niger	National Survey on Household Living Conditions	2014
Nigeria	Socio Economic Survey	2019
Republic of the Congo	Employment Survey	2009
Rwanda	Labour Force Survey	2017
Senegal	National Employment Survey	2019
Sierra Leone	Integrated Household Survey	2018

Somalia	High Frequency Survey	2017
South Africa	General Household Survey	2019
South Sudan	High Frequency Survey	2015
Tanzania	National Household Budget Survey	2012
Togo	Regional Integrated Survey on Employment and the Informal Sector	2017
Uganda	National Labour Force Survey	2017
Zambia	Labour Force Survey	2019
Zimbabwe	Labour Force Survey	2014

F. Data Appendix: Educational Attainment Data

F.1. Data Sources

Barro-Lee Database The primary data source used to measure the evolution of educational attainment is the database compiled by Barro and Lee (2013) and updates.[33] The database covers the distribution of educational attainment by age group and gender in 146 countries at five year intervals from 1950 to 2015. It covers 123 countries out of the 150 countries studied in this paper. The education categories are no schooling, incomplete primary, complete primary, incomplete secondary, complete secondary, incomplete tertiary, and complete tertiary. I interpolate linearly the share of individuals belonging to each category between missing years, and extrapolate linearly educational attainment by age and gender after 2015, so as to cover the entire 1980-2019 period.

IPUMS and Survey Data For the 27 countries absent from the Barro-Lee database, I rely on census and survey data. For Burkina Faso (1986-2006), Ethiopia (1984-2007), Guinea (1983-2014), and Palestine (1997-2017), the data source is the census microdata samples available from IPUMS International. For India, which is covered by the Barro-Lee database but displays somewhat erratic trends, I rely instead on the education modules of the national sample survey (1983-2017), which I collected and harmonized for the purpose of this paper. For the remaining 22 countries, in the absence of better data, I use cohort-level trends in educational attainment observed in the surveys collected in this paper.[34] I first aggregate the distribution of educational attainment by cohort and gender in each survey. I then derive estimates of educational attainment of the 1980 to 2019 working-age populations by taking the weighted average across cohorts belonging to the working-age population in the corresponding year.

F.2. Matching Survey and Aggregate Data

To derive accurate estimates of counterfactual income absent educational expansion, it is important to make sure that educational attainment in the survey data matches perfectly aggregate data used to derive the counterfactual. Although education levels do correlate strongly in the two sources, some inconsistencies remain. For instance, aggregate and survey

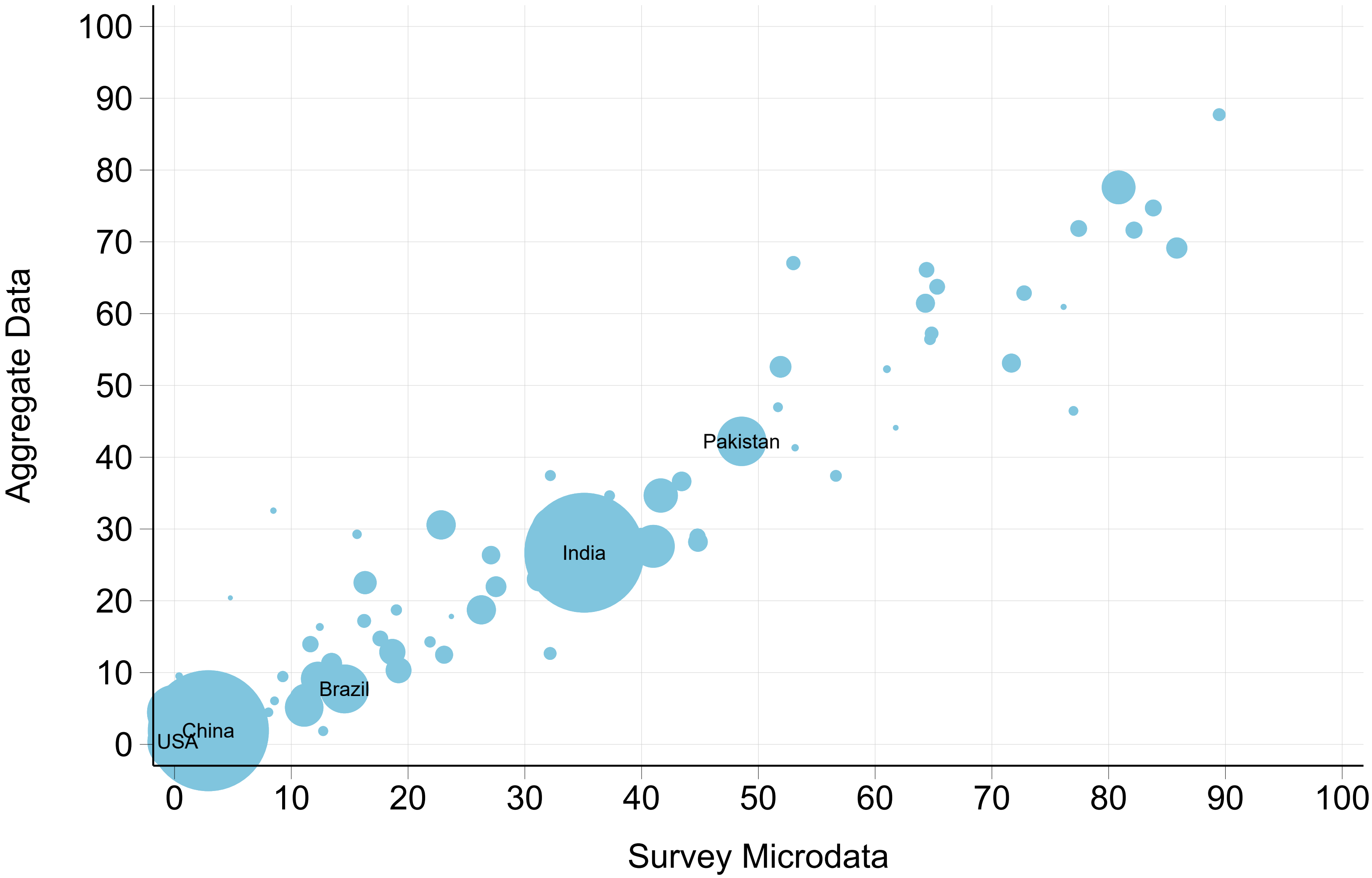
[33] See http://www.barrolee.com/.

[34] The countries are Angola, Azerbaijan, Bosnia and Herzegovina, Bhutan, Belarus, Cabo Verde, Cambodia, Chad, Djibouti, Georgia, Guinea-Bissau, Kosovo, Lebanon, Montenegro, Madagascar, Macedonia, Nigeria, Somalia, Suriname, South Sudan, Timor-Leste, and Uzbekistan.

data sometimes report incomplete degrees as complete and sometimes do not, or code lower secondary education as primary education. To make sure that the two sources coincide, I first manually recode some categories in survey and/or aggregate data, country by country, by visually inspecting the distribution of educational attainment in the two sources. The result of this manual recoding process is displayed in figures F1, F2, F3, and F4, which compare the share of the working-age population with no schooling, primary education, secondary education, and tertiary education in survey versus aggregate data. The two sources end up very close to each other after recoding.

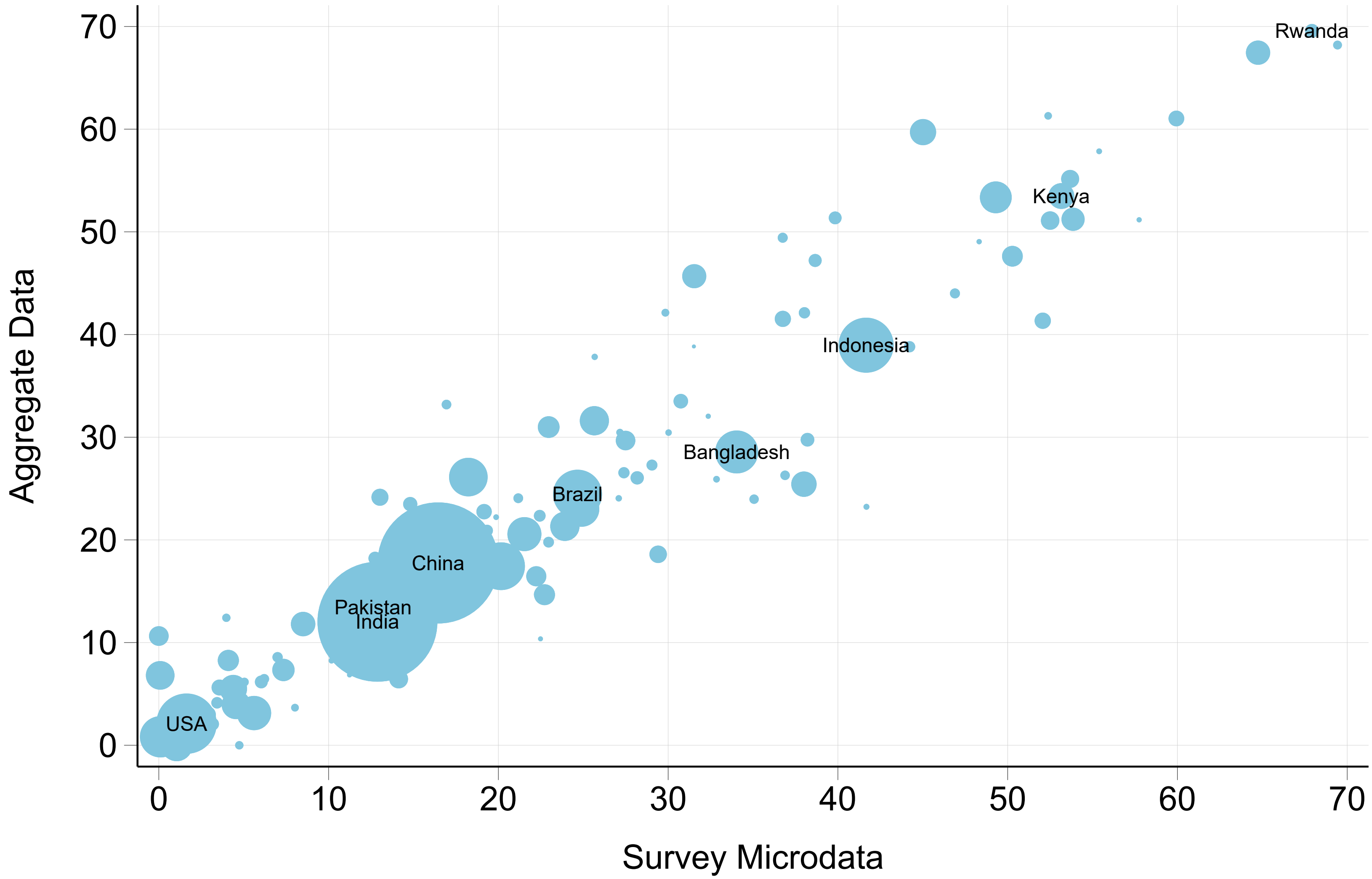
Second, I perform a final small adjustment to the sample weights of each survey to make sure that education levels by age and gender match perfectly in the two sources. I combine aggregate data on the distribution of attainment with data on total population by age and gender from the UN to derive estimates of the total number of individuals belonging to each of 40 education-age-gender cells. I then use linear calibration to ensure that total weights match the total population belonging to each cell in each country. The result is a new weight variable that ensures that the distribution of educational attainment by age and gender (and for the working-age population as a whole) in the survey data matches perfectly that observed in aggregate data.

Figure F1 – Barro-Lee Versus Survey Data: Share of Working-Age Population With No Schooling



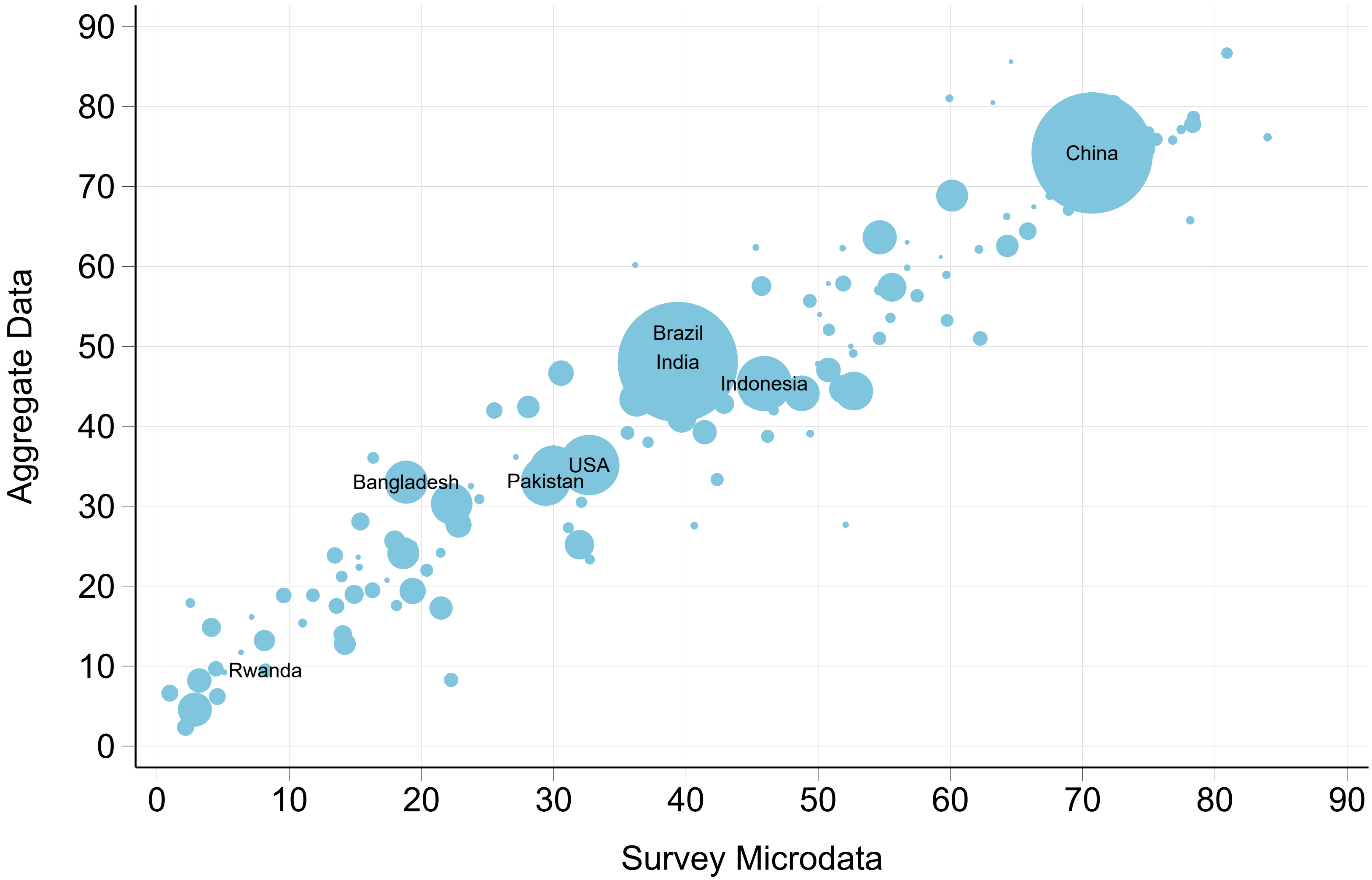
Notes. The figure compares estimates of the share of the working-age population with no schooling in the survey microdata (x-axis) and aggregate data from Barro and Lee (2013) and other sources (y-axis), after manual reclassification of educational categories in each country.

Figure F2 – Barro-Lee Versus Survey Data: Share of Working-Age Population With Primary Education



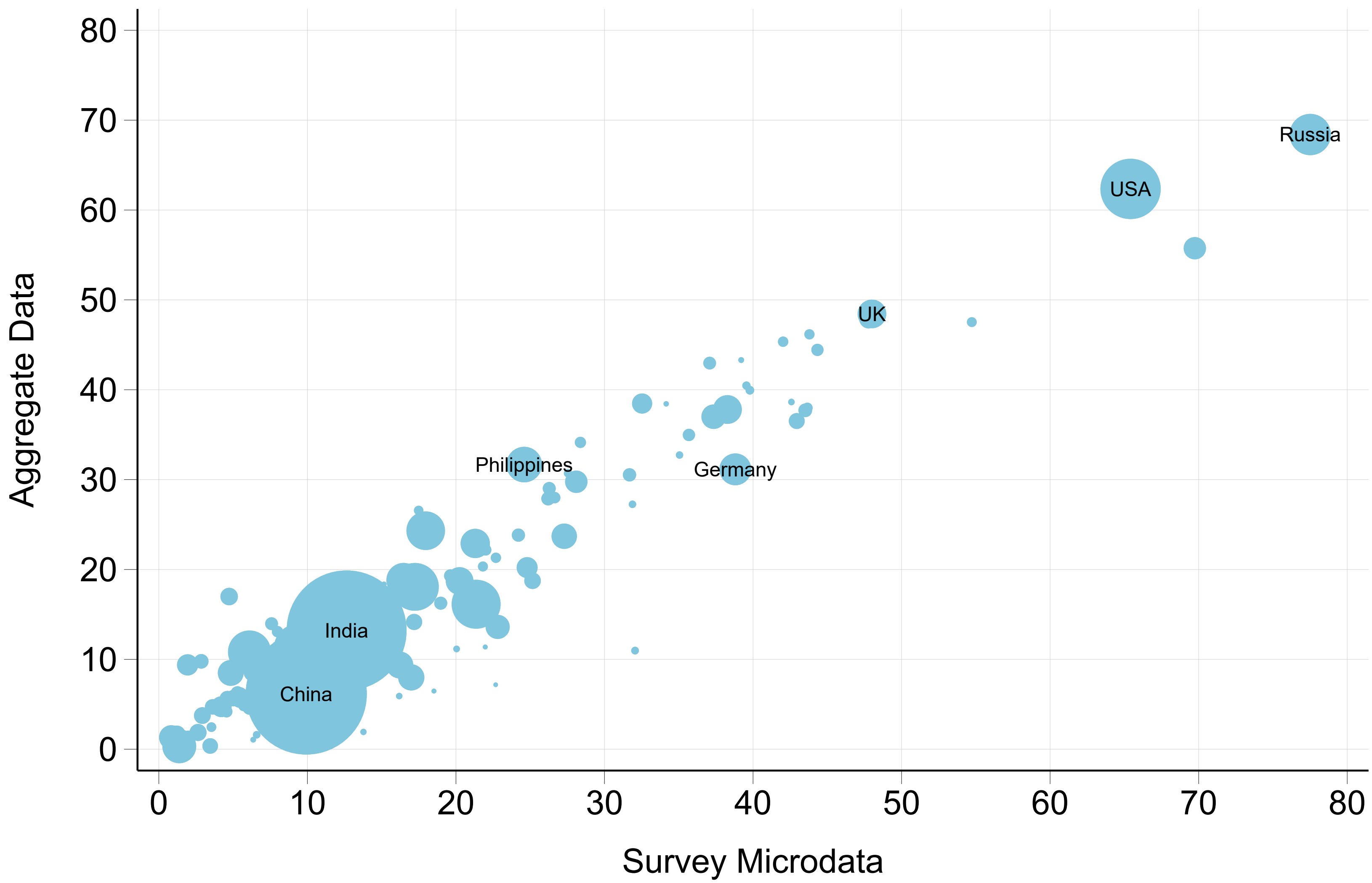
Notes. The figure compares estimates of the share of the working-age population with primary/basic education in the survey microdata (x-axis) and aggregate data from Barro and Lee (2013) and other sources (y-axis), after manual reclassification of educational categories in each country. Primary/basic education includes lower secondary education in some countries.

Figure F3 – Barro-Lee Versus Survey Data: Share of Working-Age Population With Secondary Education



Notes. The figure compares estimates of the share of the working-age population with secondary education in the survey microdata (x-axis) and aggregate data from Barro and Lee (2013) and other sources (y-axis), after manual reclassification of educational categories in each country. Secondary education excludes lower secondary education in some countries.

Figure F4 – Barro-Lee Versus Survey Data: Share of Working-Age Population With Tertiary Education



Notes. The figure compares estimates of the share of the working-age population with tertiary education in the survey microdata (x-axis) and aggregate data from Barro and Lee (2013) and other sources (y-axis), after manual reclassification of educational categories in each country.

G. Data Appendix: Returns to Schooling

G.1. OLS Estimates of Returns to Schooling

In the main analysis, I use estimates of returns to schooling by level estimated in each country. I rely on the following modified Mincerian equation:

$$\ln y_{ict} = \alpha_t + \beta_{ct}^{pri} D_{ict}^{pri} + \beta_{ct}^{sec} D_{ict}^{sec} + \beta_{ct}^{ter} D_{ict}^{ter} + X_{ict}\beta + \varepsilon_{ict} \tag{55}$$

With y_{ict} earned income of individual i in country c at time t, D_{ict}^{pri}, D_{ict}^{sec}, and D_{ict}^{ter} dummies for having reached primary, secondary, and tertiary education, and X_{ict} a vector of controls including gender, an experience quartic, and interactions between gender and the experience quartic. Earned income is the sum of all wage and self-employment income received by a given individual. I restrict the sample to all individuals aged above 15 with strictly positive income. I estimate this regression separately in each country and extract estimates of β_{ct}^{pri}, β_{ct}^{sec}, and β_{ct}^{ter}. In 47 countries with too few observations to estimate the return to primary education, I make the very conservative assumption that the return observed in 2019 is exactly zero. The same holds in 12 countries with too few observations to estimate the return to secondary education. Note that primary and secondary education do still end up having posiive effects on earnings in these countries in the benchmark specification, because imperfect substitution implies that the true return lies above the return observed in 2019 (see section 2.2).

Figure G2 plots the distribution of annualized returns to schooling by level, while figures G3 to G6 map these returns in all countries with available estimates. Average returns to a year of schooling, estimated using a Mincerian equation with individual years of schooling on the right-hand side, typically range from 3% to 20%, with a median of 9%. Returns to primary education are typically lower than returns to secondary education, which are themselves below returns to tertiary education. The return to primary education is just 5% in the median country, compared to 9% for secondary education and 13% of tertiary education.

Table G1 investigates the robustness of these results to using a standard Mincerian equation with only gender, potential experience, and potential experience squared as controls. Table G2 compares baseline estimates pooling labor and self-employment income to a specification restricting the analysis to wage income. The results are almost identical: the average return to schooling is 8.9-9.7%, while the returns to primary, secondary, and tertiary education are 4.5-4.7%, 8.4-8.8%, and 14.4-15%, respectively.

Another concern is that workers and self-employed individuals declaring positive personal income might only represent a subset of the population. This is particularly concerning in low-income countries, where a large fraction of the population often relies on subsistence agriculture and thus ends up excluded from my estimation of the returns to schooling. I investigate this concern in appendix table G3 by comparing three specifications. The first one corresponds to a standard Mincerian equation estimated at the individual level, restricting the sample to individuals declaring positive personal income. The second specification corresponds to a "household-level Mincerian equation," regressing per-capita expenditure on adults' average years of schooling. The third specification repeats the second specification, but after restricting the sample to households with at least one adult declaring positive personal income, which is useful to check whether the results are driven by selection into reporting positive income. I estimate these returns for eleven countries characterized by high poverty rates and large agricultural sectors: India, Pakistan, the Democratic Republic of the Congo, Burkina Faso, Côte d'Ivoire, Guinea-Bissau, Mali, Niger, Sénégal, and Togo. For each of these eleven countries, I was able to collect and manually harmonize survey microdata covering personal income, household expenditure, and educational attainment.

The three estimates end up falling very close to each other, amounting to a Mincerian return typically varying from 7% to 10%. Individual returns are slightly higher than household-level returns in some countries, such as India, Pakistan, and Côte d'Ivoire, which is to be expected given that variations in consumption are more driven by other factors, such as savings and transfers received by other households and the government. Yet there are also countries where individual returns are lower, such as Burkina Faso and Mali. Household-level returns before and after excluding households with no reported income are virtually identical in most countries. Together, these findings provide reassuring evidence that the returns estimated in this paper provide a good approximation of the true returns to schooling for the population as a whole.

G.2. IV Estimates of Returns to Schooling

In an alternative specification, I rely on instrumental variable estimates of the returns to schooling from a number of existing studies (see table G4), which I use to adjust Mincerian OLS returns estimated with my data. Indeed, given that IV estimates from collected studies were generally computed at a different period and using a different sample than mine, they cannot directly be used in the estimation. Another difficulty is that these returns are annualized, while the returns used in my analysis correspond to total log-point increases in earnings from reaching specific levels of educational attainment. I thus incorporate IV estimates into the estimation in two steps. First, I use the ratio of IV to OLS estimates of yearly returns to schooling from these studies to adjust yearly returns (see figure G7). Second, I exponentiate this ratio by level-specific average years of schooling to adjust total returns by level.

Formally, the total return of moving from level s_1 to level s_2 is:

$$r(s_1, s_2) = \ln(w_2) - \ln(w_1) \tag{56}$$

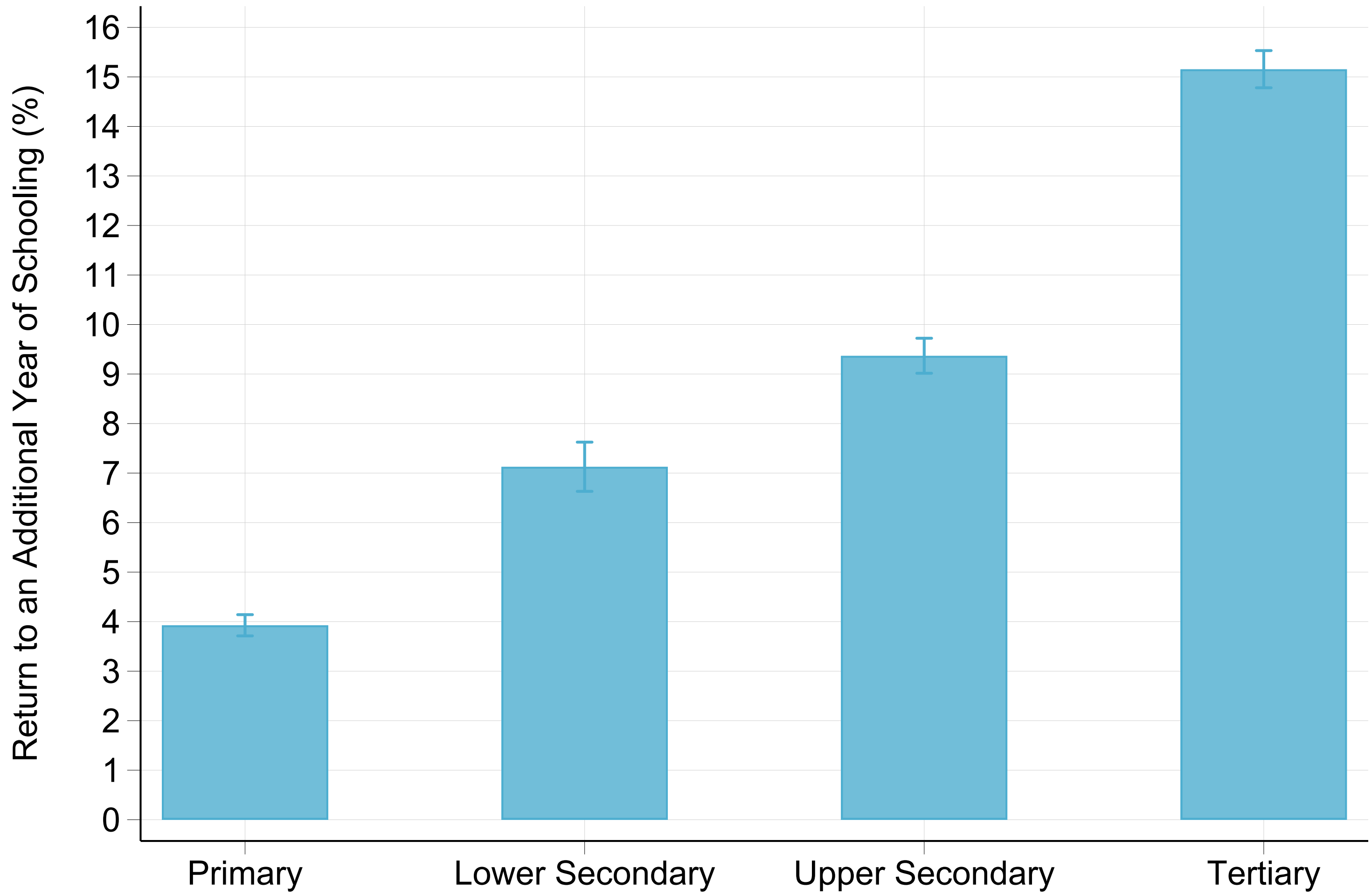
Which implies that the ratio of w_2 to w_1 is $w_2/w_1 = \exp(r(s_1, s_2))$. The corresponding annualized return to schooling is thus:

$$\beta(s_1, s_2) = \exp\left(r(s_1, s_2)\right)^{1/\mathrm{T}(s_1, s_2)} - 1 \tag{57}$$

With $\mathrm{T}(s_1, s_2)$ the difference in average years of schooling between individuals with educational attainment s_2 and s_1. We know from studies relying on quasi-experimental designs that IV estimates are higher than OLS estimates by a factor γ: $\beta^{IV}(s_1, s_2) = \gamma\beta(s_1, s_2)$. Hence, the adjusted total return to schooling is:

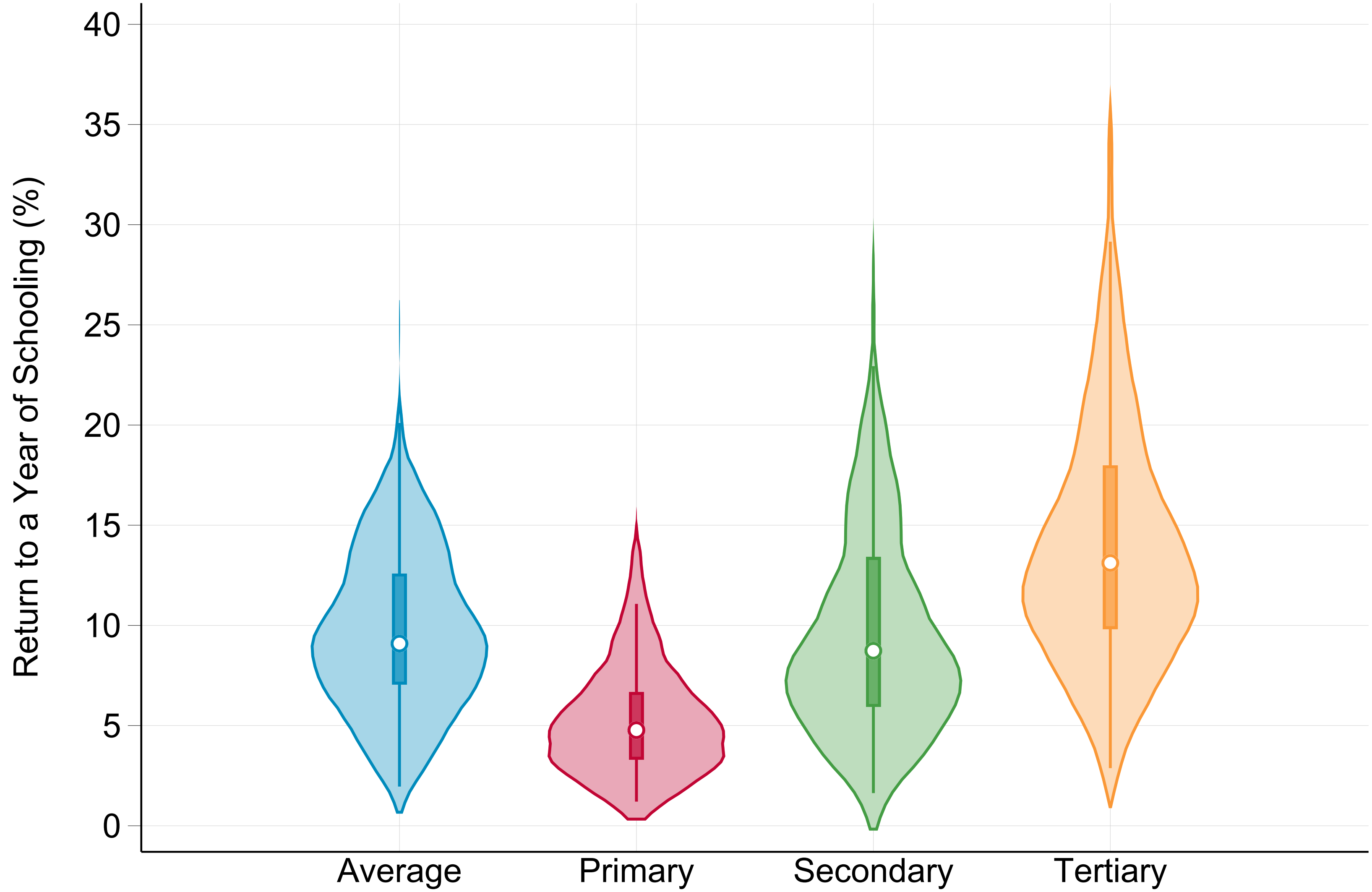
$$\begin{aligned} r^{IV}(s_1, s_2) &= \mathrm{T}(s_1, s_2) \times \ln\left(1 + \gamma\beta(s_1, s_2)\right) && (58) \\ &= \mathrm{T}(s_1, s_2) \times \ln\left[1 + \gamma\left(\exp\left(r(s_1, s_2)\right)^{1/\mathrm{T}(s_1, s_2)} - 1\right)\right] && (59) \end{aligned}$$

Figure G1 – Returns to Schooling: Pooled Estimates by Level



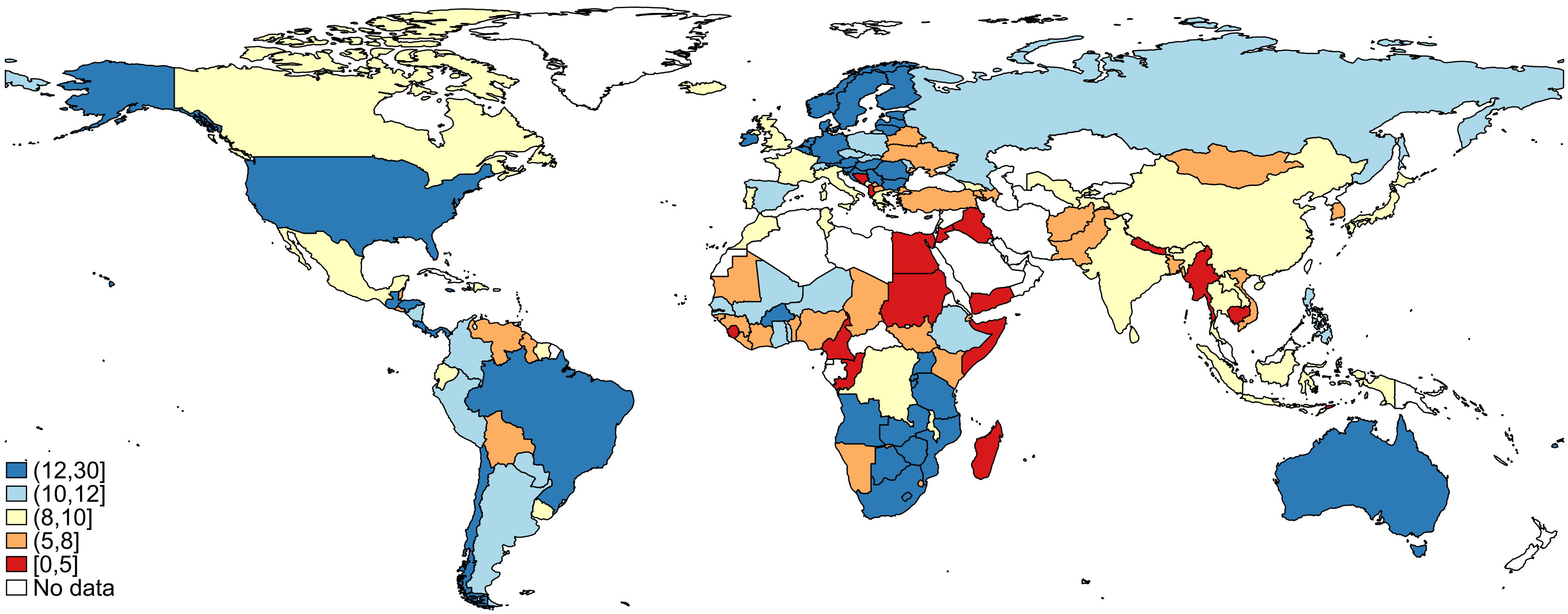
Notes. The figure reports estimates of an additional year of schooling by education level, based on a pooled regression on the full micro dataset. Primary: returns to a year of primary education. Lower secondary: return to a year of lower secondary education, restricting the sample to individuals with either primary or lower secondary education. Upper secondary: return to a year of upper secondary education, restricting the sample to individuals with either lower secondary or upper secondary education. Tertiary: return to a year of higher education, restricting the sample to individuals with either upper secondary or tertiary education. All models include controls for gender, an experience quartic, interactions between gender and the experience quartic, and country fixed effects. Observations are weighted to match each country's total population. Capped spikes represent 95% confidence intervals.

Figure G2 – Returns to Schooling: Distribution of Estimates by Level



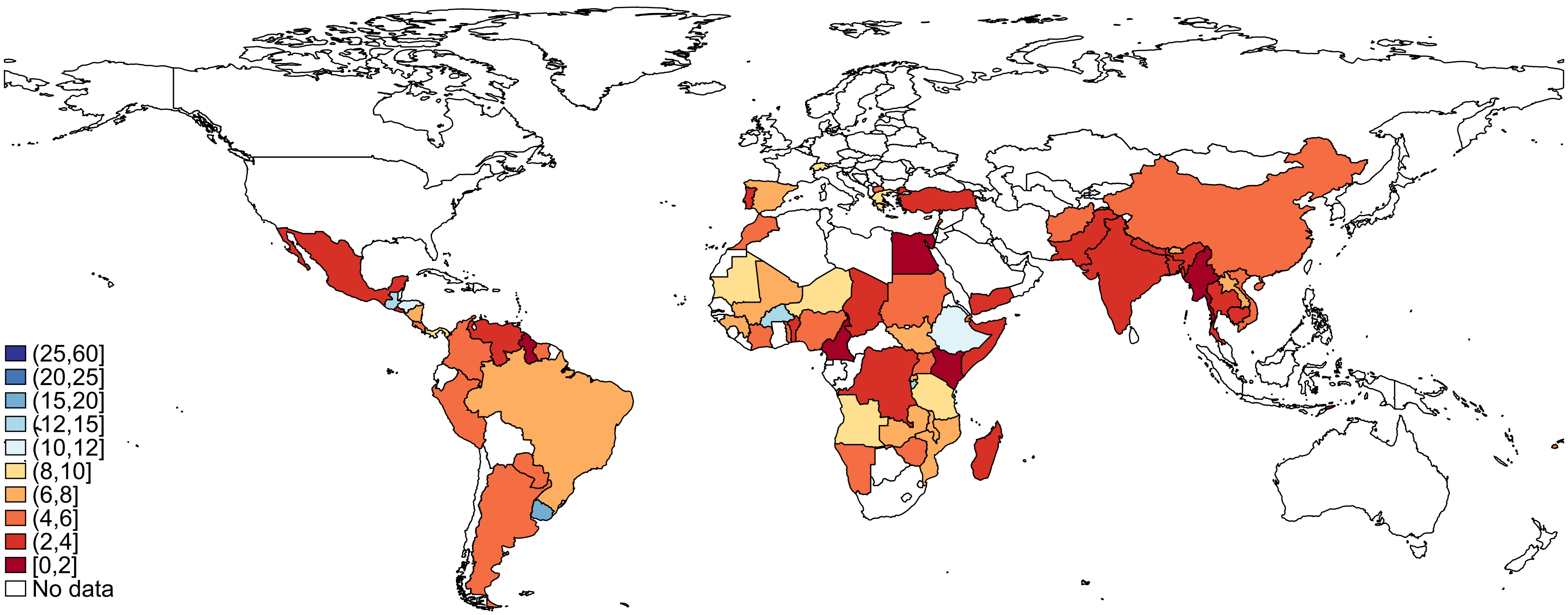
Notes. Author's computations using labor force survey microdata. The figure plots the cross-country distribution of returns to schooling by level. Estimates correspond to the effect of one additional year of schooling on the log of personal income, estimated using modified Mincerian equations controlling for an experience quartic, gender, and interactions between the experience quartic and gender. Primary: return to a year of schooling among individuals with either no schooling, some primary education, or completed primary education. Secondary: return to a year of schooling among individuals with either some primary education, completed primary education, some lower or upper secondary education, or completed upper secondary education. Tertiary: return to a year of schooling among individuals with some upper secondary education, completed upper secondary education, some tertiary education, or completed tertiary education.

Figure G3 – Return to an Additional Year of Schooling



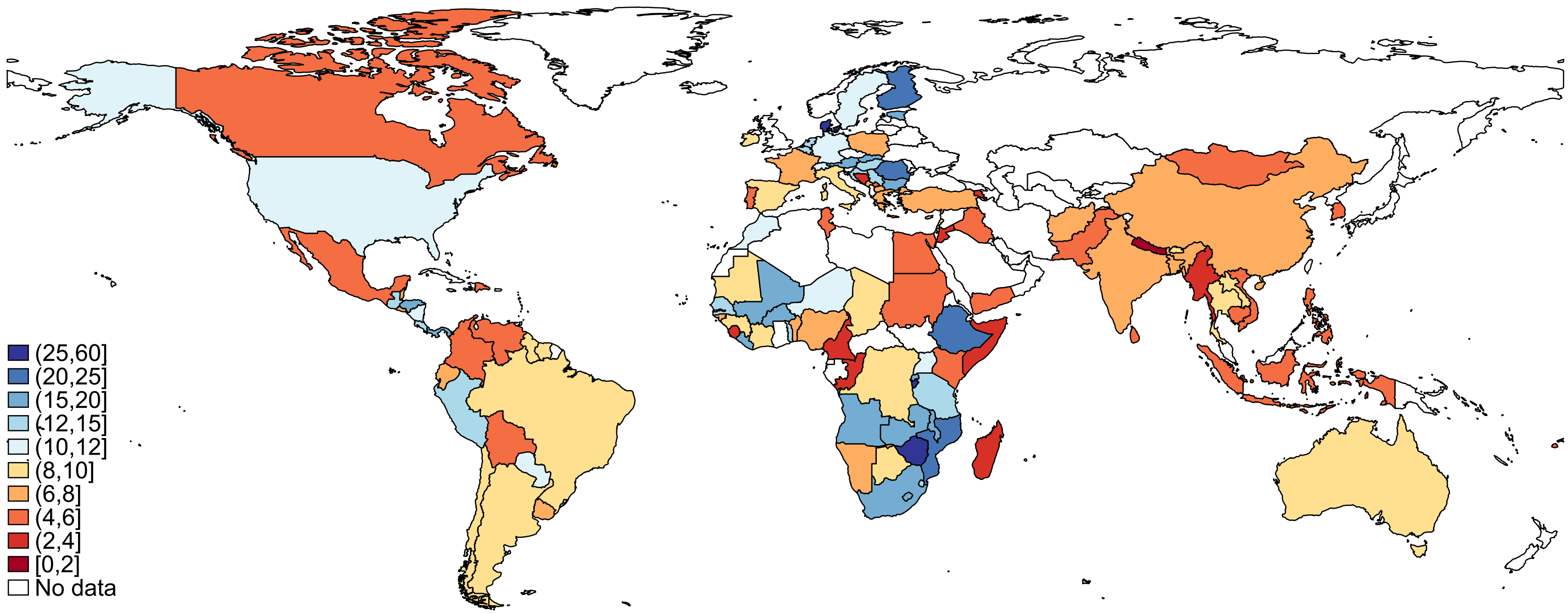
Notes.

Figure G4 – Returns to an Additional Year of Primary Education



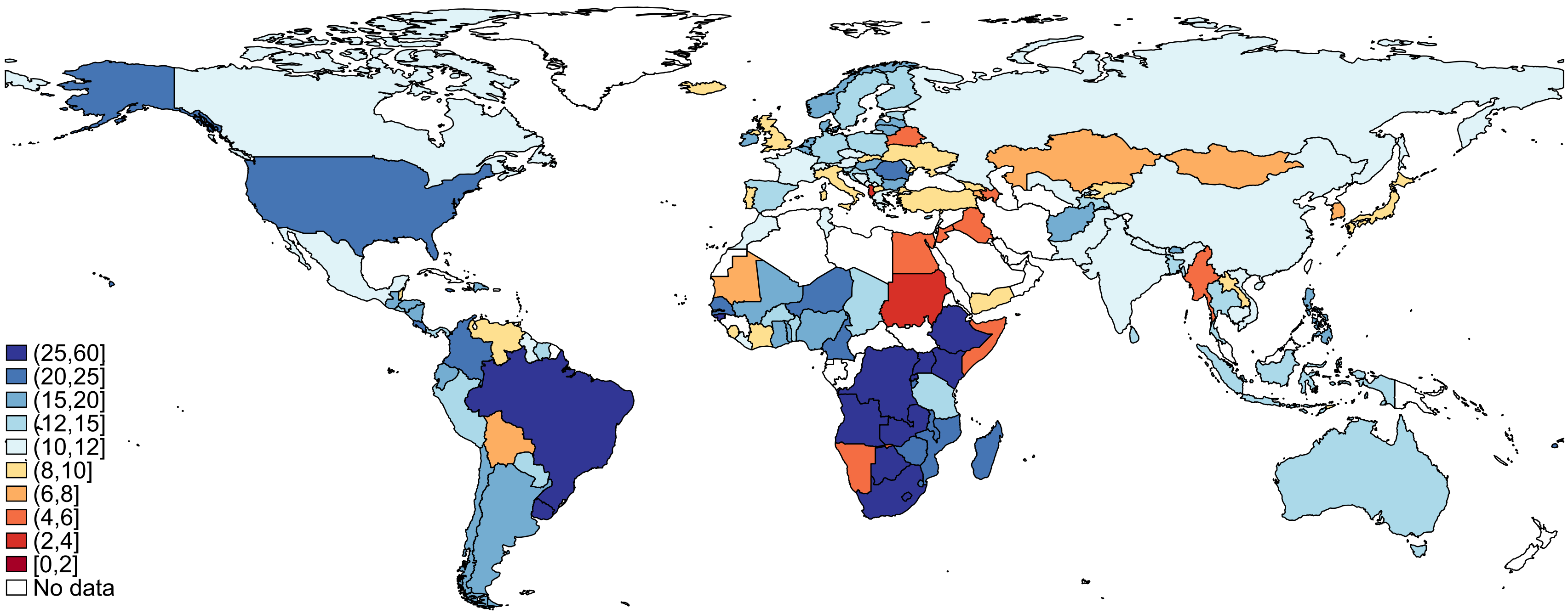
Notes.

Figure G5 – Returns to an Additional Year of Secondary Education



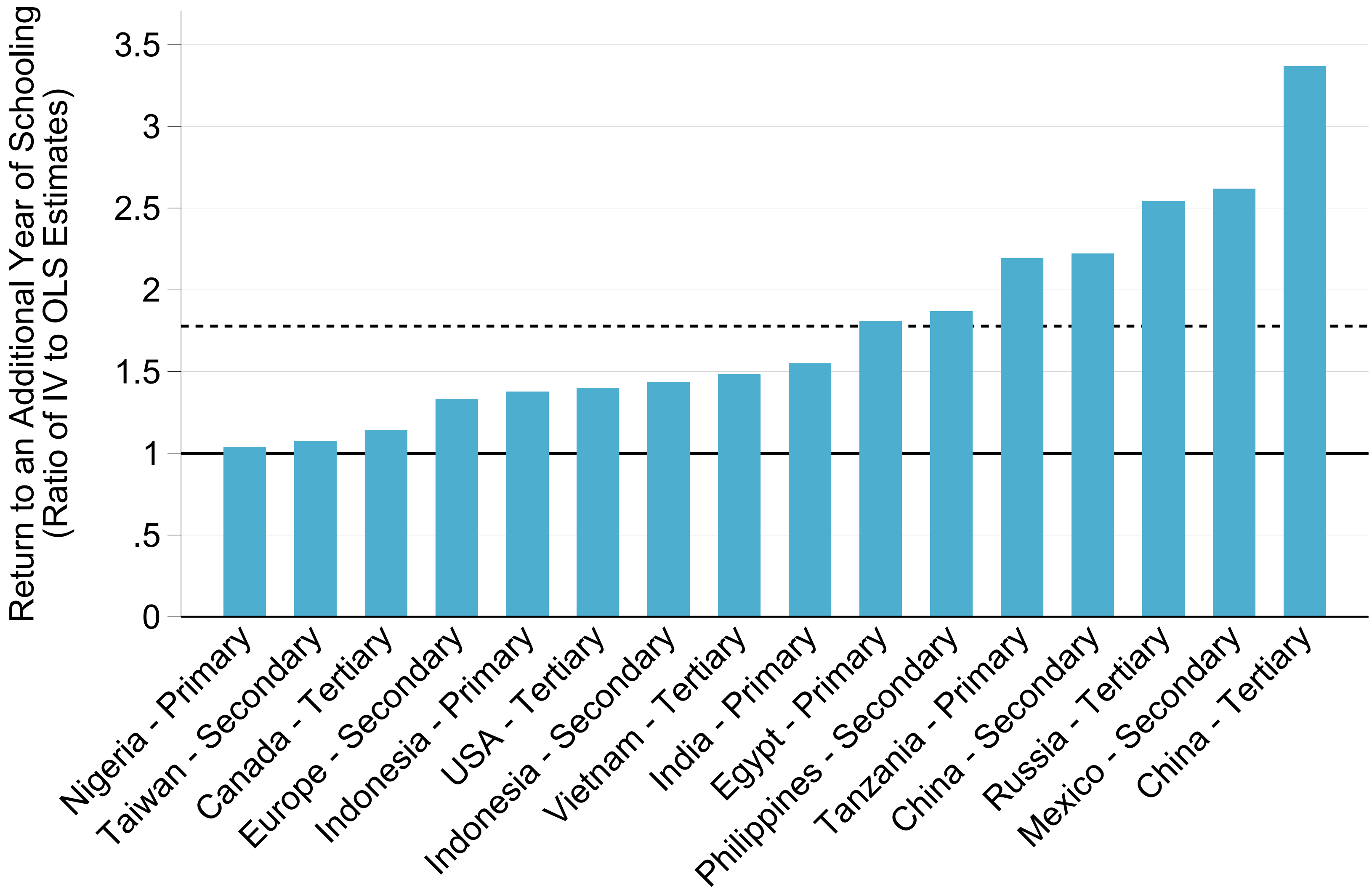
Notes.

Figure G6 – Returns to an Additional Year of Tertiary Education



Notes.

Figure G7 – Returns to Schooling: Ratio of IV to OLS Estimates



Notes. Author's elaboration compiling estimates from a number of published studies. Pri/Sec/Ter: returns to a year of primary/secondary/tertiary education. Dashed line: average ratio across all estimates.

Table G1 – Returns to Schooling: Standard versus Extended Mincer Equation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Years of Schooling	0.089***	0.090***	0.045***	0.046***	0.084***	0.085***	0.150***	0.150***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Level	All	All	Primary	Primary	Secondary	Secondary	Tertiary	Tertiary
Extended Model	No	Yes	No	Yes	No	Yes	No	Yes
N	4,912,763	4,912,763	1,493,033	1,493,033	3,184,353	3,184,353	2,772,992	2,772,992
Adj. R-squared	0.80	0.80	0.83	0.83	0.82	0.82	0.79	0.79

Notes. The table reports estimates of Mincerian returns, comparing "standard" and "extended" versions of the model by education level. Standard version: controls for gender, potential experience, and potential experience squared. Extended version: controls for gender, an experience quartic, and interactions between gender and the experience quartic, as in Lemieux (2006). Pooled regression across the full micro dataset. All estimates include country fixed effects. Observations are weighted to match each country's total population. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table G2 – Returns to Schooling: Total Personal Income Versus Wages

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Years of Schooling	0.090***	0.097***	0.046***	0.047***	0.085***	0.088***	0.150***	0.144***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Level	All	All	Primary	Primary	Secondary	Secondary	Tertiary	Tertiary
Income Concept	Total	Wages	Total	Wages	Total	Wages	Total	Wages
N	4,912,763	3,677,689	1,493,033	926,343	3,184,353	2,319,863	2,772,992	2,265,470
Adj. R-squared	0.80	0.88	0.83	0.90	0.82	0.87	0.79	0.88

Notes. The table reports estimates of Mincerian returns, comparing models including total personal income (wages + self-employment income) to models restricting the sample to wage earners. All models include controls for gender, an experience quartic, interactions between gender and the experience quartic, and country fixed effects. Observations are weighted to match each country's total population. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table G3 – Returns to Schooling: Personal Income Versus Per-Capita Consumption

	Individual Income	Consumption All Households	Consumption Households With Income Only
India	0.070***	0.063***	0.063***
	(0.001)	(0.001)	(0.001)
Pakistan	0.074***	0.071***	0.071***
	(0.001)	(0.001)	(0.001)
DR Congo	0.075***	0.071***	0.070***
	(0.002)	(0.002)	(0.002)
Burkina Faso	0.089***	0.106***	0.086***
	(0.006)	(0.003)	(0.005)
Benin	0.078***	0.068***	0.071***
	(0.003)	(0.002)	(0.003)
Côte d'Ivoire	0.073***	0.058***	0.052***
	(0.003)	(0.002)	(0.002)
Guinea-Bissau	0.041***	0.061***	0.052***
	(0.004)	(0.002)	(0.003)
Mali	0.072***	0.080***	0.061***
	(0.005)	(0.002)	(0.003)
Niger	0.108***	0.103***	0.101***
	(0.004)	(0.003)	(0.003)
Sénégal	0.078***	0.074***	0.075***
	(0.003)	(0.002)	(0.002)
Togo	0.100***	0.078***	0.076***
	(0.006)	(0.002)	(0.005)

Notes. The table compares returns to schooling estimated with three specifications. The first specification regresses individual income on individual years of schooling, controlling for age, gender, and their interaction. The second specification regresses per-capita consumption on average years of schooling of working-age adults at the household level, controlling for household size, average age, and the share of women. The third specification does the same, but after restricting the sample to households with at least one adult declaring positive personal income. India: 2019 PLFS survey. Pakistan: 2018 HIES survey. DR Congo: 2012 ECM survey. Other countries: 2018 EHCVM surveys. Robust standard errors in parenthesis. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table G4 – IV Estimates of Returns to Schooling

Source	Country	Level	OLS β	IV β	OLS SE	IV SE
Lemieux and Card (2001)	Canada	Tertiary	7	8	.2	4.4
Fang et al. (2012)	China	Secondary	9	20	.4	.6
Huang and Zhu (2022)	China	Tertiary	4.9	16.5		
Assaad et al. (2023)	Egypt	Primary	2.1	3.8	.3	4.5
Brunello, Weber, and Weiss (2015)	Europe	Secondary	4.2	5.6	.3	2.6
Khanna (2023)	India	Primary	10	15.5		
Carneiro, Lokshin, and Umapathi (2017)	Indonesia	Secondary	9	12.9	.5	4.8
Duflo (2001)	Indonesia	Primary	7.7	10.6	.06	2.2
Navarro-Sola (2021)	Mexico	Secondary	4.7	12.3	.1	1.6
Oyelere (2010)	Nigeria	Primary	2.6	2.7	.1	1.3
Sakellariou (2006)	Philippines	Secondary	6.1	11.4		
Kyui (2016)	Russian Federation	Tertiary	6.1	15.5	.25	1.1
Spohr (2003)	Taiwan	Secondary	5.4	5.8		
Delesalle (2021)	Tanzania	Primary	2.6	5.7	.01	2.1
Zimmerman (2014)	USA	Tertiary	10	14		
Vu and Vu-Thanh (2022)	Viet Nam	Tertiary	16	23.7		

Notes. The table reports instrumental variable estimates of returns to schooling from selected articles. OLS: return to schooling estimated by OLS. β: return to a year of schooling. SE: standard error associated with the estimate.